



TUGAS AKHIR - KS141501

ANALISIS DAN VISUALISASI SUARA PELANGGAN PADA PUSAT LAYANAN PELANGGAN DENGAN PEMODELAN TOPIK MENGGUNAKAN LATENT DIRICHLET ALLOCATION (LDA) STUDI KASUS: PT. PETROKIMIA GRESIK

ANALYSIS AND DATA VISUALIZATION ON CUSTOMER SERVICE APPLICATION USING TOPIC MODELLING WITH LATENT DIRICHLET ALLOCATION CASE STUDY : PT. PETROKIMIA GRESIK

**ARI AGUSTINA
NRP 5213 100 016**

**Dosen Pembimbing:
Renny Pradina K., S.T., M.T.**

**DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2017**

TUGAS AKHIR - KS141501

**ANALISIS DAN VISUALISASI SUARA PELANGGAN PADA PUSAT
LAYANAN PELANGGAN DENGAN PEMODELAN TOPIK
MENGUNAKAN LATENT DIRICHLET ALLOCATION (LDA)
STUDI KASUS: PT. PETROKIMIA GRESIK**

**ARI AGUSTINA
NRP 5213 100 016**

**Dosen Pembimbing:
Renny Pradina K., S.T., M.T.**

**DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2017**



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS141501

**ANALYSIS AND DATA VISUALIZATION ON CUSTOMER
SERVICE APPLICATION USING TOPIC MODELLING WITH
LATENT DIRICHLET ALLOCATION CASE STUDI : PT.
PETROKIMIA GRESIK**

**ARI AGUSTINA
NRP 5213 100 016**

**Supervisor:
Renny Pradina K., S.T., M.T.**

**DEPARTMENT OF INFORMATION SYSTEMS
Faculty of Information Technology
Sepuluh Nopember Institute of Technology
Surabaya 2016**

LEMBAR PENGESAHAN

ANALISIS DAN VISUALISASI SUARA PELANGGAN PADA PUSAT LAYANAN PELANGGAN DENGAN PEMODELAN TOPIK MENGGUNAKAN LATENT DIRICHLET ALLOCATION (LDA) STUDI KASUS: PT. PETROKIMIA GRESIK

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada

Jurusan Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh:

ARI AGUSTINA
5213 100 016

Surabaya, Januari 2017

**KETUA
JURUSAN SISTEM INFORMASI**

Dr. Ir. Aris Tjahyanto, M.Kom.
NIP.19650310 199102 1 001

LEMBAR PERSETUJUAN

ANALISIS DAN VISUALISASI SUARA PELANGGAN PADA PUSAT LAYANAN PELANGGAN DENGAN PEMODELAN TOPIK MENGGUNAKAN LATENT DIRICHLET ALLOCATION (LDA) STUDI KASUS: PT. PETROKIMIA GRESIK

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada

Jurusan Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh :

ARI AGUSTINA

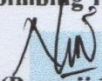
5213 100 016

Disetujui Tim Penguji : Tanggal Ujian : 11 Januari 2017
Periode Wisuda: Maret 2017

Renny Pradina K., S.T, MT


(Pembimbing 1)

Nur Aini Rakhmawati S.Kom, M.Sc.Eng, Ph.D


(Penguji 1)

Faizal Johan Atletiko, S.Kom, M.Kom


(Penguji 2)

ANALISIS DAN VISUALISASI SUARA PELANGGAN PADA PUSAT LAYANAN PELANGGAN DENGAN PEMODELAN TOPIK MENGGUNAKAN LATENT DIRICHLET ALLOCATION (LDA) STUDI KASUS: PT. PETROKIMIA GRESIK

Nama Mahasiswa : Ari Agustina
NRP : 5213 100 016
Jurusan : Sistem Informasi FTIF-ITS
Pembimbing 1 : Renny Pradina K., S.T, M.T.

ABSTRAK

Pupuk merupakan salah satu kebutuhan penting di bidang pertanian. Sebagai salah satu perusahaan yang memproduksi pupuk, PT. Petrokimia Gresik menyadari kepuasan pelanggan merupakan salah satu aspek yang menandakan keberhasilan perusahaan dalam memahami kebutuhan pelanggan.

Untuk itu, PT. Petrokimia Gresik senantiasa mengoptimalkan suara pelanggan yang masuk melalui aplikasi bernama Pusat Layanan Pelanggan. Saat ini, Bagian Promosi dan Aplikasi Produk pada PT. Petrokimia Gresik telah melakukan pencatatan terhadap seluruh suara pelanggan yang masuk melalui berbagai media, baik SMS, telepon, ataupun media massa.

Namun, dari hasil pencatatan yang telah dilakukan belum terdapat analisis lebih lanjut mengenai hal-hal apa saja yang sering dibahas oleh pelanggan saat menyampaikan suara pelanggan. Oleh karena itu, penelitian ini menganalisis topik dari suara pelanggan yang tercatat dengan pemodelan Latent Dirichlet Allocation (LDA).

Berdasarkan penelitian tersebut, pemodelan LDA telah terbukti mampu untuk mengidentifikasi topik dengan baik. Hal ini ditunjukkan dengan 35 topik yang berhasil diidentifikasi

kemudian dikelompokkan ke dalam 7 kategori. Kualitas luaran dari pemodelan topik yang dihasilkan juga cukup baik, ditunjukkan dengan nilai perplexity sebesar 34.92 dengan standar deviasi 0.49 pada 20 iterasi. Akurasi dari pengujian model yang dihasilkan adalah 83.7%. Hasil identifikasi ini, kemudian divisualisasikan dalam dashboard berbasis web.

Kata Kunci: Latent Dirichlet Allocation, Pemodelan Topik, Kepuasan Pelanggan, Dashboard

**ANALYSIS AND DATA VISUALIZATION ON
CUSTOMER SERVICE APPLICATION USING TOPIC
MODELLING WITH LATENT DIRICHLET
ALLOCATION CASE STUDY : PT. PETROKIMIA
GRESIK**

Student Name : Ari Agustina

NRP : 5213 100 016

Department : Sistem Informasi FTIF-ITS

Supervisor 1 : Renny Pradina K., S.T, M.T.

ABSTRACT

Fertilizers is one of the important needs in agriculture. PT petrokimia gresik, a fertilizer producing company, realize that customer satisfication is one of any aspect that indicates the company's success of understanding customer needs.

PT petrokimia gresik optimize their customer opinions with an application called "Pusat Layanan Pelanggan". Promotion and Application Product Function/ Department has recorded the cutomer's opinions from many media such as sms, telephone, or social media.

However, until this time, PT Petrokimia Gresik have not been analyzing their result of those records. Therefore, this research will analyze PT Petrokimia gresik customer's opinions with Latent Driecllet Allocation (LDA) model.

Based on this research, the LDA model has proved to identify the topics properly. 35 topics has successfully identified, then those topics has grouped in to 7 categories. The output wuality of this topics model is good enough, it is shown in perplexity score as big as 34.92 with deviation standard is 0.49, at 20 iteration. Accuration of the testing model is 83.7%. Result of the identification has visualized in to a dashboard in web-based.

Keywords: Latent Dirichlet Allocation, Topic Modelling, Customer Satisfaction, Dashboard

Halaman sengaja dikosongkan

KATA PENGANTAR

Puji dan syukur penulis ucapkan ke hadirat Allah SWT yang telah memberikan kekuatan dan hidayahnya kepada penulis sehingga penulis dapat menyelesaikan laporan tugas akhir dengan judul: **“Analisis dan Visualisasi Suara Pelanggan Pada Pusat Layanan Pelanggan Dengan Pemodelan Topik Menggunakan Latent Dirichlet Allocation (LDA) Studi Kasus: PT. PETROKIMIA GRESIK”** yang merupakan salah satu karya penulis untuk Jurusan Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya.

Tugas akhir ini tidak akan pernah terwujud tanpa bantuan dan dukungan dari berbagai pihak yang sudah meluangkan waktu, tenaga, dan pikirannya bagi penulis dalam menyelesaikan tugas akhir ini. Pada kesempatan ini penulis ingin mengucapkan terima kasih sebesar-besarnya kepada pihak di bawah ini, yaitu:

1. Orang tua yaitu Ibu Ngasmi dan Almarhum Bapak Pal Arifin serta keluarga yang senantiasa selalu mendoakan dan memberikan kasih sayang serta semangat tiada henti untuk menyelesaikan Tugas Akhir ini.
2. Bapak Dr. Ir. Aris Tjahyanto, M.Kom., selaku Ketua Jurusan Sistem Informasi ITS selama penulis menjalani kuliah.
3. Renny Pradina K., S.T, M.T. selaku dosen pembimbing yang telah banyak meluangkan waktu untuk membimbing, mengarahkan, dan mendukung dalam penyelesaian Tugas Akhir
4. Satya Windranuari beserta Departemen Yankomduk selaku pembimbing lapangan di PT. Petrokimia Gresik yang memberikan arahan serta mendukung terselesaikannya Tugas Akhir ini.

5. Bapak Apol Pribadi, S.T., M.T. selaku dosen wali yang senantiasa memberikan pengarahan dan motivasi selama penulis menempuh masa perkuliahan dan pengerjaan Tugas Akhir.
6. Safrina Kharisma Imandani, Kevin Setiawan, Robbigh Faubendri, Ashma Hanifah, Ervi Ritya dan Rani Oktavia yang selalu menemani dan memberikan semangat hingga penyelesaian Tugas Akhir ini.
7. Stezar Priansya, Valiant Verliando, serta teman-teman seperjuangan “ADDI Aman” yang telah membantu pengerjaan Tugas Akhir ini hingga selesai.
8. Teman-Teman BELTRANIS yang selalu memberikan semangat positif untuk menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna dan memiliki banyak kekurangan di dalamnya. Dan oleh karena itu penulis meminta maaf atas segala kesalahan yang dibuat penulis dalam buku tugas akhir ini. Penulis membuka pintu selebar-lebarnya bagi pihak yang ingin memberikan kritik dan saran, dan penelitian selanjutnya yang ingin menyempurnakan karya dari tugas akhir ini. Semoga buku tugas akhir ini bermanfaat bagi seluruh pembaca.

Surabaya, Januari 2016
Penulis

DAFTAR ISI

KATA PENGANTAR.....	ix
DAFTAR ISI.....	xvi
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvi
DAFTAR KODE.....	xviii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Tujuan	3
1.5. Manfaat.....	4
1.6. Relevansi	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Penelitian Sebelumnya	5
2.2. Dasar Teori.....	9
2.2.1. PT. Petrokimia Gresik	9
2.2.2. Topic Modelling.....	16
2.2.3. Latent Dirichlet Allocation.....	17
2.2.4. Customer Relationship Management	19
2.2.5. Visualisasi Dashboard	20
BAB III METODOLOGI	21
3.1. Tahapan Pelaksanaan Tugas Akhir	21
3.2. Uraian Metodologi	22

3.2.1.	Identifikasi Masalah	22
3.2.2.	Studi Literatur.....	22
3.2.3.	Pengumpulan dan Pemahaman Data	23
3.2.4.	Pemilihan Data	25
3.2.5.	Topic Modelling dengan Latent Dirichlet Allocation	25
3.2.6.	Analisis Topik	29
3.2.7.	Pembuatan <i>Dashboard</i>	30
BAB IV PERANCANGAN		31
4.1.	Pengambilan Data.....	31
4.2.	Seleksi Atribut.....	36
4.3.	Metodologi Implementasi Penilaian.....	39
4.3.1.	Load Data	39
4.3.2.	Pra-proses Data.....	40
4.3.3.	Proses Data	41
4.3.4.	Analisis Topik	41
4.3.5.	Klasifikasi Data	42
4.3.6.	Konstruksi Perangkat Lunak	42
BAB V IMPLEMENTASI		49
5.1.	Perangkat Penelitian	49
5.2.	Memuat Data	49
5.3.	Pengerjaan Pra-Proses Data.....	50
5.3.1.	Stemming.....	50
5.3.2.	Pendefinisian Stopword.....	51

5.3.3.	Tokenization	53
5.4.	Pengerjaan Proses Data	54
5.4.1.	Penentuan Iterasi	55
5.4.2.	Penentuan Jumlah Topik	56
5.5.	Analisis Topik	57
5.6.	Klasifikasi Data	57
5.5.	Konektor PHP dengan Python.....	58
5.6.	Pengerjaan Visualisasi Data	59
BAB VI HASIL DAN PEMBAHASAN.....		71
6.1.	Load Data	71
6.2.	Percobaan pra-proses data	72
6.3.	Pembuatan Model LDA	74
6.3.1.	Penentuan jumlah iterasi	74
6.3.2.	Penentuan jumlah topik	80
6.4.	Analisis Topik	92
6.5.	Klasifikasi Data	103
6.6.	Pengujian Model pada Aplikasi	110
BAB VII KESIMPULAN DAN SARAN		111
7.1.	Kesimpulan.....	111
7.2.	Saran.....	112
REFERENSI		113
BIODATA PENULIS		115
LAMPIRAN A		117
LAMPIRAN B		132

Halaman sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2.1. Alur penyampaian suara pelanggan melalui PLP	11
Gambar 2.2. Tampilan Aplikasi PLP	12
Gambar 2.3. Form Lembar Pantau Petrokimia Gresik	13
Gambar 2.4. Form Pengelolaan Suara Pelanggan di Aplikasi	15
Gambar 2.5. Alur mengenerate dokumen	16
Gambar 2.6. Konsep Topic Modelling [9]	17
Gambar 2.7. Metode LDA [11]	18
Gambar 2.8. Distribusi Topik LDA [9]	19
Gambar 3. 1. Tahapan Pelaksanaan Tugas Akhir	21
Gambar 3. 2. Langkah-langkah melakukan LDA [6]	26
Gambar 4.1. Alur Loading Data	40
Gambar 4.2. Alur Pra-proses Data	40
Gambar 4.3. Alur Pemrosesan Data	41
Gambar 4.4. Alur Analisis Topik	42
Gambar 4.5. Perancangan Database Aplikasi Pusat Layanan Pelanggan	43
Gambar 4.6. Alur menghubungkan antarmuka dengan model python	44
Gambar 4.7. Jumlah suara pelanggan dan jumlah pelanggan	45
Gambar 4.8. Distribusi Topik Tahunan	45
Gambar 4.9. Detail distribusi topik berdasarkan kategori	46
Gambar 5. 1. Form Suara Pelanggan	61
Gambar 5.2. Halaman Utama Aplikasi	62
Gambar 5.3. Detail Distribusi per Kategori Topik	63
Gambar 5.4. Kabar Terbaru Layanan Suara Pelanggan	64
Gambar 5. 5. Detail Kabar Terbaru Layanan Suara Pelanggan	65
Gambar 5. 6. Edit Topik Kabar Terbaru	66
Gambar 5. 7. Edit Topik Kabar Terbaru Setelah Pengisian Form Suara Pelanggan	66

Gambar 5.8. Distribusi Topik Berdasarkan Kategori per Bulan	67
Gambar 5.9. Distribusi Detail Topik per Kategori dan Bulan	68
Gambar 5.10. Total Distribusi Topik Tiap Kategori Berdasarkan Bulan	69
Gambar 6. 1. Percobaan 1 Evaluasi Perplexity dengan 30 iterasi.....	77
Gambar 6. 2. Percobaan 2 Evaluasi Perplexity dengan 30 iterasi	78
Gambar 6. 3. Percobaan 3 Evaluasi Perplexity dengan 30 iterasi	79
Gambar 6. 4. Rata-Rata Perplexity 50 Topik	82
Gambar 6. 5. Standar Deviasi Perplexity 50 Topik.....	91
Gambar 6. 6. Kategori Topik.....	95
Gambar 6. 7. Kategori Topik Produk	97
Gambar 6. 8. Kategori Topik Harga	98
Gambar 6. 9. Kategori Topik Pembelian.....	98
Gambar 6. 10. Kategori Topik Stock/Distribusi.....	99
Gambar 6. 11. Kategori Topik Promosi	100
Gambar 6. 12. Kategori Topik Pelayanan	101
Gambar 6. 13. Kategori Topik Lainnya.....	102
Gambar 6. 14. Distribusi topik setiap ID.....	104
Gambar 6. 15. Distribusi Tiap Topik.....	105
Gambar 6. 16. Distribusi Topik per Kategori.....	107
Gambar 6. 17. Distribusi Topik Tiap Kategori per Tahun ...	108
Gambar 6. 18. Total Distribusi Topik Tiap Kategori per Bulan	109

DAFTAR TABEL

Tabel 2.1. Penelitian Sebelumnya	5
Tabel 3.2. Case Folding.....	27
Tabel 3.3. Hasil Stemming pada data	27

Tabel 3.4. Stopword removal pada data	28
Tabel 3.5. Tokenization pada data.....	28
Tabel 3.6. Tabel hasil pembentukan korpus	29
Tabel 4.1. Keterangan Atribut Database	31
Tabel 4.2. Contoh Data Suara Pelanggan.....	35
Tabel 4.3. Hasil Seleksi Atribut Data.....	37
Tabel 4.4. Hasil Seleksi Atribut	38
Tabel 4.5. Kabar Terbaru Suara Pelanggan.....	47
Tabel 4.6. Kabar terbaru suara pelanggan kategori X	47
Tabel 6. 1. Jumlah data suara pelanggan.....	71
Tabel 6. 2. Kata-kata pada topik 30	83
Tabel 6. 3. Kata-kata pada topik 35	85
Tabel 6. 4. Kata-kata pada topik 40	87
Tabel A. 1. Perplexity Percobaan 1 Tanpa Stemming.....	117
Tabel A. 2. Perplexity Percobaan 2 Tanpa Stemming.....	118
Tabel A. 3. Perplexity Percobaan 3 Tanpa Stemming.....	119
Tabel A. 4. Perplexity Percobaan 1 dengan Stemming	120
Tabel A. 5. Perplexity Percobaan 2 dengan Stemming	122
Tabel A. 6. Perplexity Percobaan 3 dengan Stemming	123
Tabel A. 7. Perplexity Percobaan 1 dengan Stemming dan Pengecekan.....	124
Tabel A. 8. Perplexity Percobaan 2 dengan Stemming dan Pengecekan.....	125
Tabel A. 9. Perplexity Percobaan 3 dengan Stemming dan Pengecekan.....	126
Tabel A. 10. Hasil Uji Coba Model.....	128
Tabel B.1. Hasil Pengujian Data	133
Tabel B.2. Hasil Klasifikasi Jumlah Dokumen berdasarkan ID Topik	141
Tabel B.3. Hasil Klasifikasi Data berdasarkan Topik	142
Tabel B.4. Hasil Klasifikasi Data berdasarkan Kategori.....	143

Tabel B.5. Hasil Klasifikasi Data berdasarkan Kategori Tiap Tahun.....	143
Tabel B. 6. Hasil Klasifikasi Data per Kategori Berdasarkan Bulan	144

DAFTAR KODE

Kode 5.1. Loading Data.....	50
Kode 5.2. Stemming pada data suara pelanggan	51
Kode 5.3. Stopword Removal	52
Kode 5.4. Tokenization pada suara pelanggan	53
Kode 5.5. Pembuatan Dictionary.....	54
Kode 5.6. Print dictionary	54
Kode 5.7. Print jumlah kata yang terdapat dalam dictionary	55
Kode 5.8. Membuat Korpus	55
Kode 5.9. Percobaan 30 iterasi	56
Kode 5.10. LDA Model dengan jumlah topik 3	57
Kode 5.11. Print hasil topik dan distribusi kata.....	57
Kode 5.12. Klasifikasi topik	58
Kode 5.13. Penghubung model python dengan kode PHP.....	59

BAB I

PENDAHULUAN

Pada bab pendahuluan akan diuraikan proses identifikasi masalah penelitian yang meliputi latar belakang masalah, perumusan masalah, batasan masalah, tujuan tugas akhir, manfaat kegiatan tugas akhir dan relevansi terhadap pengerjaan tugas akhir. Berdasarkan uraian pada bab ini, harapannya gambaran umum permasalahan dan pemecahan masalah pada tugas akhir dapat dipahami.

1.1. Latar Belakang

PT Petrokimia Gresik, merupakan salah satu perusahaan pupuk terbesar di Asia Tenggara. Sebagai anak perusahaan yang tergabung dalam Pupuk Indonesia Holding Company, Petrokimia Gresik berusaha memberikan yang terbaik bagi setiap konsumennya. Saat ini, PT Petrokimia Gresik mampu memproduksi lebih dari 4 juta ton pupuk dalam setiap tahunnya. Pupuk yang dihasilkan pun beraneka ragam, mulai dari Urea, NPK, ZA, Pupuk Organik, dan lain sebagainya. Keberadaan PT Petrokimia Gresik adalah untuk mendukung program Pemerintah dalam rangka meningkatkan produksi pertanian dan ketahanan pangan Nasional[1].

Sebagai perusahaan industri yang mengutamakan kepuasan pelanggan, PT. Petrokimia Gresik berusaha untuk mengetahui dan memahami kebutuhan pelanggan yaitu melalui layanan suara pelanggan dan riset pasar. Melalui layanan suara pelanggan, petani seringkali mengeluhkan berbagai permasalahan yang cukup beragam, mulai dari info produk hingga cara penggunaan pupuk. Dari beragam keluhan yang disampaikan pelanggan tersebut, perlu ditindaklanjuti agar kepuasan pelanggan dapat memiliki hasil yang lebih optimal. Menjaga kepuasan pelanggan dengan layanan yang diberikan dapat meningkatkan keuntungan perusahaan. Selain itu, dengan menjaga kepuasan pelanggan, akan membuat

pelanggan memiliki sifat yang loyal terhadap perusahaan, sehingga tidak mudah berpindah terhadap produk atau layanan yang dimiliki oleh kompetitor[2].

Saat ini, PT Petrokimia Gresik telah melakukan pencatatan dari layanan suara pelanggan melalui tiga tahapan proses, yaitu pencatatan manual dengan menggunakan form suara pelanggan, kemudian menyalin form tersebut ke dalam penyimpanan database berbasis Microsoft Excel, dan menginputkan kembali data yang telah disimpan pada Microsoft Excel ke dalam aplikasi Pusat Layanan Pelanggan. Namun, belum ada analisis lebih lanjut untuk memahami apa saja yang dikeluhkan. Pemahaman ini akan memberikan kemampuan untuk menanggapi suara pelanggan secara lebih bermakna dan pada level yang lebih strategis.

Oleh karena itulah, penelitian ini akan menawarkan solusi dalam melakukan analisis *topic modelling* terhadap perilaku pelanggan yang dikeluhkan melalui suara pelanggan. Analisis *topic modelling* digunakan untuk mengetahui topik-topik apa saja yang sering muncul, sehingga memudahkan Bagian Promosi dan Aplikasi Produk dalam melakukan mapping jawaban yang dihasilkan, serta strategi yang dapat diterapkan selanjutnya. Penelitian ini menggunakan metode Latent Dirichlet Allocation atau yang selanjutnya disebut dengan LDA. Metode LDA merupakan sebuah metode yang tergolong sebagai salah satu bentuk dari text mining untuk menemukan pola tertentu pada sebuah dokumen. LDA membuat dokumen menghasilkan beberapa macam topik yang berbeda, sehingga tidak secara spesifik mengelompokkan dokumen kedalam sebuah topik tertentu.

Hasil dari analisis *topic modelling* menggunakan LDA akan ditampilkan dengan sebuah visualisasi berupa *dashboard* pada website yang dapat menampilkan trend dari topik-topik yang sering dibahas, sebagai suatu sistem yang dapat membantu pengambilan keputusan.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan permasalahan yang menjadi fokus dan akan diselesaikan dalam tugas akhir ini antara lain:

1. Bagaimana melakukan topic modelling untuk menganalisis topik-topik apa saja yang sering dikeluhkan pelanggan dari layanan suara pelanggan PT Petrokimia Gresik menggunakan metode Latent Dirichlet Allocation?
2. Bagaimana menampilkan hasil analisis untuk topik yang sering dibahas dalam layanan suara pelanggan sehingga dapat membantu pengambilan keputusan pada Bagian Promosi dan Aplikasi Produk PT Petrokimia Gresik?

1.3. Batasan Masalah

Dari permasalahan yang disebutkan di atas, batasan masalah dalam tugas akhir ini adalah:

1. Data yang digunakan adalah data pusat layanan pelanggan PT Petrokimia Gresik tahun 2014-2016
2. Terdapat kebijakan/landasan dalam menentukan penggolongan topik dari suara pelanggan
3. Pembuatan aplikasi adalah sebatas dashboard dari hasil analisis *topic modelling* layanan suara pelanggan

1.4. Tujuan

Berdasarkan latar belakang dan permasalahan yang telah dijelaskan, maka tujuan dari tugas akhir ini adalah sebagai berikut:

1. Melakukan analisis topic modelling untuk mengetahui topik-topik yang sering dibahas pelanggan dari layanan suara pelanggan PT Petrokimia Gresik menggunakan metode Latent Dirichlet Allocation.
2. Menampilkan hasil analisis topik yang sering dibahas dalam layanan suara pelanggan sehingga dapat membantu pengambilan keputusan pada Bagian Promosi dan Aplikasi Produk PT Petrokimia Gresik.

1.5. Manfaat

Melalui tugas akhir ini diharapkan dapat memberikan manfaat, yaitu:

Bagi Penulis :

Mendapatkan pengetahuan mengenai cara melakukan analisis terhadap topik-topik yang sering dibicarakan oleh pelanggan dengan menggunakan topic modelling untuk mengetahui hubungan antara hal-hal yang dibahas sehingga menjadi topik dengan apa yang dapat dilakukan oleh perusahaan.

Bagi Instansi:

PT Petrokimia Gresik dapat terbantu untuk melakukan pengolahan data pelanggan untuk mengetahui perilaku pelanggan khususnya dalam layanan suara pelanggan mengenai hal-hal yang sering dibahas, sehingga dari hasil pengolahan data dengan menggunakan topic modelling dapat memberikan pengetahuan bagi PT Petrokimia Gresik untuk menindaklanjuti dengan strategi promosi.

Bagi Masyarakat:

Masyarakat dapat terbantu untuk mendapatkan informasi sesuai dengan yang mereka harapkan terhadap PT. Petrokimia Gresik.

1.6. Relevansi

Tugas akhir ini berkaitan dengan mata kuliah Statistika, Sistem Cerdas, Konstruksi Perangkat Lunak, Pemrograman berbasis Web, Penggalian Data dan Analitika Bisnis, Sistem Pendukung Keputusan, dan Kecerdasan Bisnis

BAB II

TINJAUAN PUSTAKA

Bab ini akan menjelaskan mengenai penelitian sebelumnya dan dasar teori yang dijadikan acuan atau landasan dalam pengerjaan tugas akhir ini. Landasan teori akan memberikan gambaran secara umum dari landasan penjabaran tugas akhir ini.

2.1. Penelitian Sebelumnya

Bagian ini akan menjelaskan mengenai penelitian sebelumnya yang telah dilakukan serta dijadikan acuan pada penelitian ini. Tabel 2.1. menampilkan daftar penelitian sebelumnya yang mendasari tugas akhir ini:

Tabel 2.1. Penelitian Sebelumnya

Judul penelitian	Metode	Penulis	Hasil yang Didapatkan
<i>Topic detection and tracking for conversational content by using conceptual dynamic latent Dirichlet Allocation</i>	<i>Conceptual Dynamic Latent Dirichlet Allocation</i>	Jui-Feng Yeh, Yi-Shan Tan, Chen-Hsien Lee	Paper pada penelitian ini fokus terhadap deteksi topik serta pelacakan pada konten percakapan. Konten percakapan menunjukkan struktur topik yang lebih rapuh daripada ilmu pengetahuan atau artikel berita lainnya. Hal ini disebabkan karena kata-kata yang berlebihan yang diucapkan saat

			<p>percakapan seringkali membingungkan untuk melakukan pendeteksian topik. Metode LDA menunjukkan kinerja yang lebih rendah karena memiliki keterbatasan terhadap penafsiran kemunculan kata-kata. Sedangkan metode CDLDA memiliki hasil yang lebih signifikan karena menganggap informasi yang disampaikan juga didasarkan terhadap cara berbicara, konsep semantik serta kesamaan arti atau sinonim. Kesamaan arti mempunyai struktur yang lebih kompleks terhadap penafsiran kata. Hasil penelitian membuktikan bahwa pendekatan yang dilakukan praktis serta efisien untuk melakukan pelacakan dan deteksi topik.[3]</p>
--	--	--	--

Identifikasi Topik Informasi Publik Media Sosial di Kota Surabaya Berdasarkan Klasterisasi Teks pada Twitter dengan Menggunakan Algoritma K-Means	Klasterisasi dengan Algoritma K-Means	Moh. Hasan Basri	Penelitian yang dilakukan dapat menemukan topik-topik yang sering dibahas pada sosial media kota Surabaya dengan menggunakan beberapa metode, diantaranya SVM, dengan hasil yang memiliki prosentase paling tinggi adalah SVM kernel “linear”[4].
<i>Advanced topic modeling for social business intelligence</i>	<i>Topic Modelling</i>	Enrico Gallinucci, Matteo Golfarelli, Stefano Rizzi	Penelitian ini membahas mengenai SBI atau <i>Social Business Intelligence</i> sebagai bidang yang relevan untuk dilakukan penelitian dengan deteksi topik hirarkis berdasarkan ketidakteraturan dan kekentalan dari hirarki, integrasi dengan hirarki bisnis serta agregasi berbasis semantik. Penelitian ini memberikan saran bahwa agar dapat menentukan topik yang sering

			berubah dengan menggunakan <i>topic hierarchy generations</i> , sehingga topik yang dihasilkan dapat dilakukan secara otomatis dan membuat hirarki topik tersebut menjadi meta-star[5].
<i>Topic Models As A Novel Approach To Identify Themes In Content Analysis: The Example Of Organizational Research Methods</i>	<i>Latent Dirichlet Allocation (LDA)</i>	Ajai Gaur	Penelitian ini membahas mengenai topic modelling yang dilakukan oleh ORM dengan 15 topik yang dihasilkan dengan prosetase keseragaman distribusi pada setiap topik yaitu 6,66%. Hal ini menunjukkan bahwa pada setiap topik dengan kemunculan diatas 6,66% lebih sering muncul daripada distribusi yang seragam dan yang dibawah 6,66% tidak ditampilkan.[6]

2.2. Dasar Teori

Bagian ini akan menjelaskan mengenai dasar teori yang dijadikan acuan pada penelitian ini.

2.2.1. PT. Petrokimia Gresik

PT Petrokimia Gresik merupakan produsen pupuk terlengkap di Indonesia, yang pada awal berdirinya disebut Proyek Petrokimia Surabaya. Saat ini, perusahaan menempati lahan seluas 450 hektar berlokasi di Kabupaten Gresik, Propinsi Jawa Timur. PT Petrokimia Gresik memproduksi berbagai macam pupuk, seperti: Urea, Za, SP-36, ZK, NPK Phonska, NPK Kebomas, dan pupuk organik Petroganik. PT Petrokimia Gresik merupakan anak perusahaan dari Pupuk Indonesia Holding Company yang menunjang kebijaksanaan dan program pemerintah di bidang ekonomi dan pembangunan nasional pada umumnya, khususnya di bidang industri, perdagangan dan angkutan. Adapun visi misi PT Petrokimia Gresik adalah:

Visi

Menjadi produsen pupuk dan produk kimia lainnya yang berdaya saing tinggi dan produknya paling diminati konsumen.

Misi

- Mendukung penyediaan pupuk nasional untuk tercapainya program swasembada pangan.
- Meningkatkan hasil usaha untuk menunjang kelancaran kegiatan operasional dan pengembangan usaha perusahaan.
- Mengembangkan potensi usaha untuk mendukung industri kimia nasional dan berperan aktif dalam community development.

Budaya Perusahaan

- Safety (Keselamatan) – Mengutamakan keselamatan dan kesehatan kerja serta pelestarian lingkungan hidup dalam setiap kegiatan operasional.
- Innovation (Inovasi) – Meningkatkan inovasi untuk memenangkan bisnis.
- Integrity (Integritas) – Mengutamakan integritas di atas hal.

- Synergistic Team (Tim yang Sinergis) – Berupaya membangun semangat kelompok yang sinergistik.
- Customer Satisfaction (Kepuasan Pelanggan) – Memanfaatkan profesionalisme untuk peningkatan kepuasan pelanggan.

2.2.1.1. Kondisi Kekinian Bagian Promosi dan Aplikasi Produk

Bagian Promosi dan Aplikasi Produk pada PT Petrokimia Gresik merupakan sebuah bagian dibawah naungan Departemen Pelayan Komunikasi Produk atau biasa disebut dengan Yankomduk. Departemen Yankomduk, dibawah oleh GM Pemasaran, dibawah Direktur Komersil. Tugas utama dari Bagian Promosi dan Aplikasi Produk adalah diantaranya untuk meningkatkan penjualan produk, mengenalkan produk baru, mempertahankan dan meningkatkan brand image dan perusahaan, serta memberikan kepuasan pelanggan melalui layanan informasi. Saat ini, bentuk kegiatan promosi yang dilakukan adalah :

1. Demplot
2. Sosialisasi
3. Publikasi
4. Sponsorship
5. Temu Pelanggan
6. Pameran
7. Promosi Penjualan
8. Pusat Layanan Pelanggan

Kegiatan promosi ini dilakukan rutin setiap bulan secara bergantian dengan tujuan yang berbeda-beda dengan indikator keberhasilan yang berbeda pula. Selain dengan melakukan kegiatan secara pro aktif, Departemen Yankomduk juga mewadahi komunikasi dengan pelanggan melalui aplikasi bernama Pusat Layanan Pelanggan atau PLP. Aplikasi PLP ini ditujukan sebagai pusat sarana informasi dan komunikasi dengan konsumen, menampung keluhan/permasalahan pelanggan, serta menjalin

komunikasi timbal balik dengan konsumen. Alur dari kondisi penyampaian suara pelanggan ditampilkan pada Gambar 2.1.



Gambar 2.1. Alur penyampaian suara pelanggan melalui PLP

Penyampaian komunikasi dari pelanggan dapat berupa telepon bebas pulsa, sms, faximile maupun email. Departemen mendefinisikan keberhasilan dari Pusat Layanan Pelanggan jika permintaan informasi konsumen dapat terlayani dengan memuaskan dan keluhan atau complain dapat diselesaikan dengan baik bersama unit terkait.

2.2.1.2. Proses Pengisian Pusat Layanan Pelanggan

Aplikasi Pusat Layanan Pelanggan dimaksudkan untuk memudahkan Bagian Promosi dan Aplikasi Produk dalam melakukan rekap hasil dari suara pelanggan yang disampaikan dalam bentuk keluhan maupun pertanyaan. Detail aplikasi dari PLP, ditampilkan pada Gambar 2.2.

PT PETROKUMIA GRESEK
Petrol Layan Pelanggan

Login as: Admin
[Logout]

Depan Respon Pelanggan Knowledgebase Master Laporan

Daftar Respon Pelanggan

9-2016 | 9-2016 | Set

Informasi Normal Baru Filter

Menampilkan 115 respon pelanggan

Ada 0 Laporan Harian Produksi Petrogrik


Tampilkan

Cari

No.	Tanggal	Reflog	P	L	Respon Pelanggan	Status	Nama	Akare	Kategori	Media
1	26-09-2016	P.0914			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Ery Dwiyanto	0856-2323430	Pertanyaan	SMS
2	26-09-2016	P.0913			Konsumen menanyakan cara pembelian Kaptan.	Close	Joko Susanto / CV. Dinasti Cabe Jemu	0852-33107281	Pertanyaan	Telepon
3	26-09-2016	P.0912			Konsumen menanyakan harga pupuk NPK Kebomas.	Close	Mugh	0812-51033381	Pertanyaan	Telepon
4	26-09-2016	P.0911			Konsumen menanyakan cara pembelian pupuk Phoska.	Close	Sarjono	0853-87669066	Pertanyaan	SMS
5	26-09-2016	P.0910			Konsumen menanyakan cara untuk melakukan pembelian Petro Fish dan organik nontsubidi.	Close	Miranto / PT. Gunung Payung Pakis Raya	0813-44838313	Pertanyaan	SMS
6	26-09-2016	P.0909			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Hilis	0857-25235884	Pertanyaan	Telepon
7	25-09-2016	P.0908			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Sholeh Agung Nugraha	0815-6732132	Pertanyaan	SMS
8	25-09-2016	P.0907			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Fahyudianto	0818-07138986	Pertanyaan	SMS
9	25-09-2016	P.0906			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Triana	0858-64477745	Pertanyaan	SMS
10	25-09-2016	P.0905			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Nugraha	0812-30498138	Pertanyaan	SMS
11	23-09-2016	P.0904			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Dian	0852-25104731	Pertanyaan	Telepon
12	23-09-2016	P.0903			Konsumen menanyakan tentang recruitment karyawan PT Petrokumia Gresik.	Close	Supriyanto	0852-13878763	Pertanyaan	Telepon

Gambar 2.2. Tampilan Aplikasi PLP

Aplikasi PLP digunakan sejak tahun 2010. Semua hasil dari keluhan suara pelanggan, baik melalui media sms, telepon, ataupun media cetak, dicatat ke dalam aplikasi. Sebelum melakukan proses pencatatan ke dalam aplikasi, PT. Petrokimia Gresik memiliki beberapa alur pemrosesan informasi mulai dari penulisan secara manual pada form lembar pantau seperti pada Gambar 2.3.

Pusat Layanan Pelanggan PT Petrokimia Gresik  PTROKIMIA GRESIK Form Pengelolaan Suara Pelanggan FM-30-0112		Jenis Pemasalahan : <input type="checkbox"/> Produk <input type="checkbox"/> Harga <input type="checkbox"/> Pembelian <input type="checkbox"/> Stock/Distribusi <input type="checkbox"/> Promosi <input type="checkbox"/> Pelayanan Lainnya	Media Akses : <input type="checkbox"/> Telepon <input type="checkbox"/> SMS <input type="checkbox"/> Email <input type="checkbox"/> Surat <input type="checkbox"/> Facsimile <input type="checkbox"/> Sales Spv.
Level Suara Pelanggan <input type="checkbox"/> Urgent <input type="checkbox"/> Normal		Dibujukan kepada unit kerja Pusat Layanan Pelanggan	
Nama/Instansi		Unit kerja penerima	
Alamat		# Nomor Registrasi	
Nomor Telepon		Pekerjaan	
Jenis Suara Pelanggan <input type="checkbox"/> Informasi <input type="checkbox"/> Pertanyaan <input type="checkbox"/> Keluhan		Email	
Segmen Pelanggan Substansi <input type="checkbox"/> RO/KK <input type="checkbox"/> Kelompok Tani <input type="checkbox"/> Alokasi pupuk bersubsidi		Non Substansi <input type="checkbox"/> Nomor DO..... <input type="checkbox"/> Tanggal pembelian..... Non Pupuk <input type="checkbox"/> Nomor DO..... <input type="checkbox"/> Tanggal pembelian.....	
Nama Produk :		Nama Produk :	
Kelompok Pelanggan <input type="checkbox"/> Corporate <input type="checkbox"/> Retail		Nama Produk :	
Media dalam mendapatkan informasi mengenai perusahaan			
Media elektronik <input type="checkbox"/> Televisi <input type="checkbox"/> Radio	Media cetak <input type="checkbox"/> Koran <input type="checkbox"/> Tabloid <input type="checkbox"/> Majalah	Media online <input type="checkbox"/> Website resmi perusahaan <input type="checkbox"/> Facebook <input type="checkbox"/> Twitter	Media Lainnya <input type="checkbox"/> Petugas lapangan/OS <input type="checkbox"/> Distributor/Kios <input type="checkbox"/> Media luar ruang
Nama Media dan Program:	Nama Media dan Rubrik:	Alamat Media:	Keterangan:
Mohon untuk mengisi suara pelanggan dalam kolom-kolom dibawah ini secara detail :			
Deskripsi		Suara pelanggan diterima oleh PLP	
		Tanggal <input type="text"/>	
		Paraf <input type="text"/>	
		Nama <input type="text"/>	
		Tlp/Ext. <input type="text"/> 2830	
		Opening Manager Promocans	

Gambar 2.3. Form Lembar Pantau Petrokimia Gresik

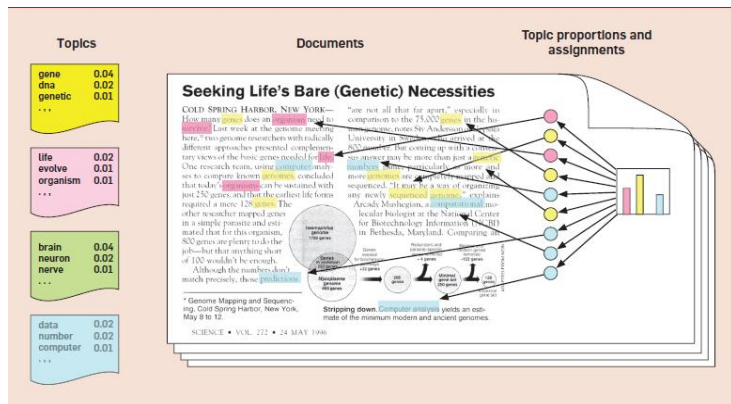
Setelah semua informasi dari pelanggan ditulis pada *form* Pengelolaan Suara Pelanggan, selanjutnya dilakukan pemindahan data dari *form* tersebut ke dalam sebuah database berbasis Microsoft Excel. Dari data yang ada pada Microsoft Excel tersebut kemudian dilakukan proses penginputan kembali ke dalam aplikasi PLP. Proses yang saat ini berjalan adalah melalui 3 kali tahap penginputan dengan jenis database yang berbeda-beda. Hal ini disebabkan karena aplikasi yang saat ini berjalan, tidak memiliki database yang dapat dikelola dan diakses dengan baik. Selain itu, *form* yang ditampilkan pada aplikasi mengenai suara pelanggan yang ditampilkan pada Gambar 2.4, tidak merepresentasikan *form* yang saat ini digunakan karena kehilangan beberapa *field* informasi yang disebabkan oleh perbaikan dari hasil *assessment* yang dilakukan terhadap perusahaan.

Gambar 2.4. Form Pengelolaan Suara Pelanggan di Aplikasi

2.2.1.3. Assessment di PT. Petrokimia Gresik

Assessment merupakan sebuah tahap atau proses yang wajib dilalui oleh perusahaan untuk melaporkan hasil kinerja yang telah dilakukan oleh perusahaan selama ini. *Assessment* merupakan tahap penilaian atau audit internal terkait dengan perkembangan setiap departemen. PT. Petrokimia Gresik adalah salah satu perusahaan yang wajib melakukan *assessment* terkait dengan kinerja yang dilakukan, terutama pada induk perusahaan yaitu Pupuk Indonesia Holding Company. Selama proses *assessment*, PT. Petrokimia Gresik melaporkan terkait tugas dan fokus utama dari setiap departemen, seperti contohnya Bagian Promosi dan Aplikasi Produk, maka melaporkan apa saja yang terkait dengan suara pelanggan, baik dari jumlah suara pelanggan yang masuk hingga media akses yang digunakan pelanggan untuk menyampaikan aspirasinya. Saat ini, *assessment* yang dilakukan pada PT.

dalam setiap dokumen. Algoritma ini, dapat digunakan untuk pengembangan dalam pencarian, ataupun meringkas teks yang ada dalam dokumen[7]. *Topic Modelling* dapat mengatur kumpulan kata berdasarkan tema yang ditemukan. Selain itu, *topic modelling* juga dapat diaplikasikan untuk berbagai jenis data, seperti yang saat ini dilakukan adalah untuk mencari pola pada data genetik, gambar, ataupun pada sosial media. *Topic Modelling* merupakan salah satu bentuk dari *text-mining* yang merupakan sebuah metode untuk menemukan dan melacak kelompok kata dalam kumpulan dokumen [8]. Konsep *topic modelling* menurut Blei, ditunjukkan pada Gambar 2.6.



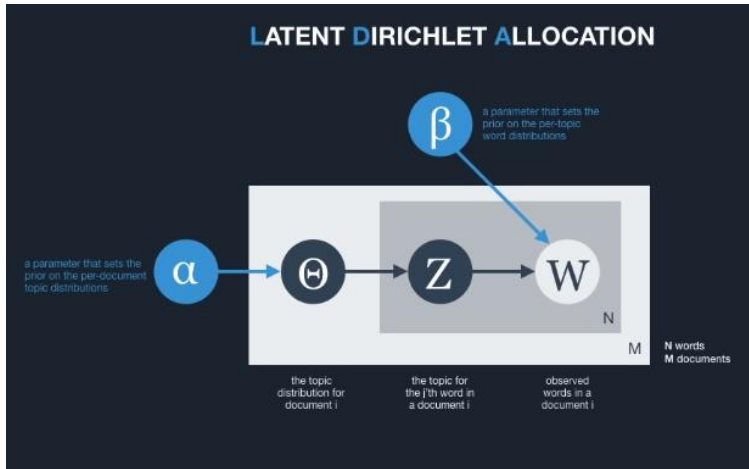
Gambar 2.6. Konsep Topic Modelling [9]

Blei, menjelaskan bahwa dokumen mengandung proporsi tersendiri dari topik-topik yang dibahas pada sebuah dokumen.

2.2.3. Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) merupakan sebuah metode statistika yang digunakan sebagai model untuk menganalisis suatu dokumen. LDA berusaha untuk melihat dokumen dengan cara mundur untuk menemukan satu set topik yang mungkin telah dikoleksi. LDA merepresentasikan dokumen dengan berbagai topik yang dibuat berdasarkan probabilitas tertentu [10]. Probabilitas topik, merepresentasikan kejelasan

dari suatu dokumen [9]. Blei menjelaskan bahwa dalam sebuah dokumen terdapat beberapa komponen, yang ditunjukkan pada Gambar 2.7.

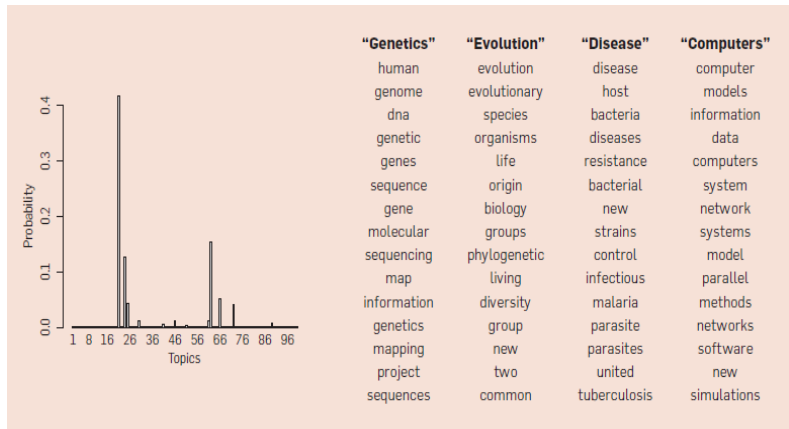


Gambar 2.7. Metode LDA [11]

Alpha menggambarkan sebuah parameter yang digunakan untuk menghitung bagaimana distribusi topik dalam dokumen. Semakin besar nilai alpha yang dimiliki oleh suatu dokumen, menandakan semakin banyaknya campuran topik yang dibahas dalam dokumen. Semakin rendah nilai alpha menunjukkan bahwa dokumen hanya membahas sedikit topik tertentu. Sedangkan Z merepresentasikan topik dari kata tertentu pada sebuah dokumen. Beta merupakan parameter yang digunakan untuk menghitung distribusi kata dalam topik. Semakin tinggi nilai beta, maka semakin banyak kata-kata yang ada di dalam topik. Semakin kecil nilai beta, maka semakin sedikit kata-kata yang ada di dalam topik sehingga lebih spesifik.

Ide dasar dari LDA adalah bahwa dalam dokumen, merepresentasikan campuran topik secara acak, dimana setiap topik digolongkan berdasarkan distribusi antar kata. Sebagai salah satu contoh dari Blei, distribusi topik yang ditampilkan

dengan kumpulan kata-kata pada dokumen ditunjukkan dengan Gambar 2.8.



Gambar 2.8. Distribusi Topik LDA [9]

LDA tidak hanya digunakan untuk melakukan pendeteksian topik saja, namun LDA juga digunakan sebagai salah satu tools untuk melakukan analisis *Business Intelligence* pada Bank, yaitu untuk mengetahui hubungan antara kebijakan tertentu dengan trend yang dihasilkan[12]. Selain pada industri perbankan, LDA juga sering digunakan untuk berbagai penelitian lain seperti pada konten percakapan [3], bahkan hingga data *software engineering*[13].

2.2.4. Customer Relationship Management

Customer Relationship Management merupakan suatu cara yang dilakukan oleh perusahaan untuk membangun hubungan dengan target pelanggan untuk memotivasi pelanggan dalam melakukan pembelian pada sebuah organisasi saat ini maupun masa depan serta memotivasi yang lainnya untuk turut serta membeli[2]. Membangun hubungan dengan pelanggan dapat diawali dengan memahami kebutuhan pelanggan, memberikan perhatian terhadap apa yang dikeluhkan pelanggan, memberikan layanan terbaik untuk pelanggan, serta

meningkatkan indeks kepuasan pelanggan. Konsep tingkat kepuasan pelanggan merupakan suatu dasar yang harus dimiliki dalam pekerjaan dan organisasi[14]. Sebagai hasil dari hubungan yang dibangun antara perusahaan dengan pelanggan, adalah tingkat loyalitas pelanggan terhadap perusahaan naik.

Menjaga hubungan pelanggan dengan baik, adalah cara yang paling efektif untuk menjadi perusahaan yang permanen dan menguntungkan[15]. Menurut konsep dasar dari hubungan pelanggan, setidaknya, jika mencari pelanggan baru adalah sesuatu yang sulit dilakukan, menjaga hubungan pelanggan dapat menjadi solusi karena pelanggan lama dapat menarik pelanggan-pelanggan baru karena kepercayaan terhadap perusahaan.

2.2.5. Visualisasi Dashboard

Salah satu cara yang bisa dilakukan untuk melakukan *monitoring* atau pemantauan terhadap perusahaan, termasuk *marketing* adalah sebuah *dashboard*. Dengan adanya *dashboard marketing*, maka perusahaan dapat melakukan efisiensi terhadap biaya pengeluaran yang cukup besar, ketepatan *marketing* yang dilakukan, serta penilaian kinerja dari apa yang telah dilakukan oleh perusahaan, khususnya di bidang *marketing*[16]. *Dashboard* dapat menjadi salah satu solusi untuk menghilangkan *knowledge gap* yang sering terjadi. Saat ini, *dashboard* tidak hanya digunakan sebagai satu-satunya *tools* yang ada pada suatu departemen, namun *dashboard* merupakan suatu hal yang umum dilakukan pada seluruh departemen pada perusahaan, seperti *dashbord sales* ataupun *dashboard produksi*. *Dashboard* digunakan untuk melihat performa perusahaan berdasarkan KPI

BAB III METODOLOGI

Bagian ini menjelaskan metodologi yang digunakan dalam pengerjaan tugas akhir ini. Metodologi ini diperlukan sebagai panduan secara sistematis dalam pengerjaan tugas akhir.

3.1. Tahapan Pelaksanaan Tugas Akhir

Pada sub bab ini akan menjelaskan mengenai metodologi dalam pengerjaan tugas akhir. Metodologi dapat dilihat pada.

Input	Proses	Output
- Issue permasalahan - Proses Bisnis Yankomduk	Identifikasi Masalah	- Solusi Permasalahan - Topik Tugas Akhir
- Permasalahan - Jurnal Ilmiah	Studi Literatur	Metode text mining yang cocok untuk jenis data
Solusi Permasalahan	Pengumpulan dan Pemahaman Data	Data Pusat Layanan Pelanggan
Data Pusat Layanan Pelanggan	Pemilihan Data	Data siap olah
Data siap olah	<i>Topic Modelling dengan Latent Dirichlet Allocation</i>	Topik hasil LDA
Topik hasil LDA	Analisis Topik	Kategorisasi Topik
Kategorisasi Topik	Pembuatan <i>Dashboard</i>	Visualisasi kemunculan kategori topik

Gambar 3. 1. Tahapan Pelaksanaan Tugas Akhir

3.2.Uraian Metodologi

Berdasarkan metodologi penelitian di atas, penjelasan setiap tahap akan dijelaskan sebagai berikut.

3.2.1. Identifikasi Masalah

Pada tahap ini dilakukan identifikasi masalah terkait dengan perusahaan yang digunakan sebagai studi kasus, yaitu PT. Petrokimia Gresik. Identifikasi masalah juga mencakup mengenai pemahaman proses bisnis dari PT. Petrokimia Gresik. Identifikasi ini akan memberikan dua hasil, yaitu permasalahan yang ditemui, serta metode yang tepat digunakan untuk menganalisis permasalahan tersebut. Permasalahan yang ditemui adalah belum adanya analisis lebih lanjut mengenai apa saja yang sering dibahas oleh pelanggan. Hal-hal yang sering dibahas oleh pelanggan di PT Petrokimia Gresik disampaikan melalui berbagai media, baik media cetak maupun sms atau telepon yang kemudian dicatat melalui Form Suara Pelanggan. Proses yang dilakukan oleh PT Petrokimia Gresik melalui 3 tahapan yang cukup panjang, dimulai dari pencatatan dengan media konvensional yaitu kertas, kemudian dilakukan pencatatan melalui database sederhana yaitu Microsoft Excel, dan terakhir disalin kembali untuk dimasukkan ke dalam aplikasi Pusat Layanan Pelanggan.

3.2.2. Studi Literatur

Pada tahap ini dilakukan studi literatur sesuai dengan permasalahan yang telah diidentifikasi, yaitu mengenai hal yang sering dibahas pada layanan suara pelanggan PT Petrokimia Gresik, konsep CRM dalam memahami kebutuhan pelanggan, serta pembuatan *dashboard* yang mendukung pengetahuan tentang kebutuhan pelanggan. Literatur disini adalah penjelasan konsep-konsep atau penelitian sebelumnya yang pernah dilakukan terkait dengan permasalahan. Literatur berasal dari dokumentasi dalam bentuk buku, jurnal, maupun website. Untuk dapat mengetahui analisis apa saja yang sering dibahas oleh pelanggan dari ketiga proses yang telah

dijelaskan pada tahapan sebelumnya, diperlukan sebuah metode yang dapat menggolongkan dokumen suara pelanggan ke dalam beberapa topik bahasan yaitu menggunakan pemodelan topik dengan LDA. Tujuan dari tahapan ini adalah dapat memahami konsep, metode, dan teknologi dengan menggunakan *Latent Dirichlet Allocation* (LDA) untuk dapat memberi solusi mengenai permasalahan yang akan digunakan dalam penyusunan tugas akhir.

3.2.3. Pengumpulan dan Pemahaman Data

Pada tahap ini dilakukan pengumpulan data dan informasi yang diperlukan untuk melakukan penyelesaian terhadap permasalahan. Data diperoleh dari pencatatan yang dilakukan dari *form* suara pelanggan, yaitu berupa database dengan format xls dari Bagian Promosi dan Aplikasi Produk PT. Petrokimia Gresik untuk mengetahui hal-hal apa saja yang dibahas secara lengkap dari permasalahan yang akan diselesaikan. Data yang dikumpulkan, dimulai dari bulan Januari 2014 hingga Juli 2016. *Fields* dari data adalah Jenis Masukan, Perihal, Deskripsi, Dari, Akses Media, Tanggal, Analisa Penyebab Potensial, Tindak Lanjut, Pencegahan Masalah, Ditindak Lanjuti Oleh, Tanggal, serta Status. Contoh data yang didapatkan dari Bagian Promosi dan Aplikasi Produk PT. Petrokimia Gresik ditampilkan pada Tabel 3.1.

Tabel 3.1. Data Suara Pelanggan

No	Jenis Masukan	Perihal	Deskripsi	Dari	Akses	Media	Tanggal	Analisa penyebab potensial	Tindak Lanjut	Pencegahan Masalah	Ditindak Lanjuti Oleh	Tanggal	Status
	Pertanyaan	Pembelian	Konsumen menanyakan cara untuk melakukan pembelian pupuk KCl.	xxxxxxxxx xxxxxxxxx	0812-xxxxxx	SMS	01/07/16	Konsumen tertarik untuk menggunakan dan mencoba manfaat produk KCl.	Penjelasan ke konsumen, pembelian pupuk KCl harus atas nama instansi (PT/CV) dan minimal order 30 ton. Harga pupuk KCl Rp 5.100/kg, harga tersebut dapat berubah tergantung kondisi pasar.	Konsumen diarahkan untuk melihat persyaratan pembelian produk pupuk nonsubsidi di website resmi PT Petrokimia Gresik.	PLP	01/07/16	Close

3.2.4. Pemilihan Data

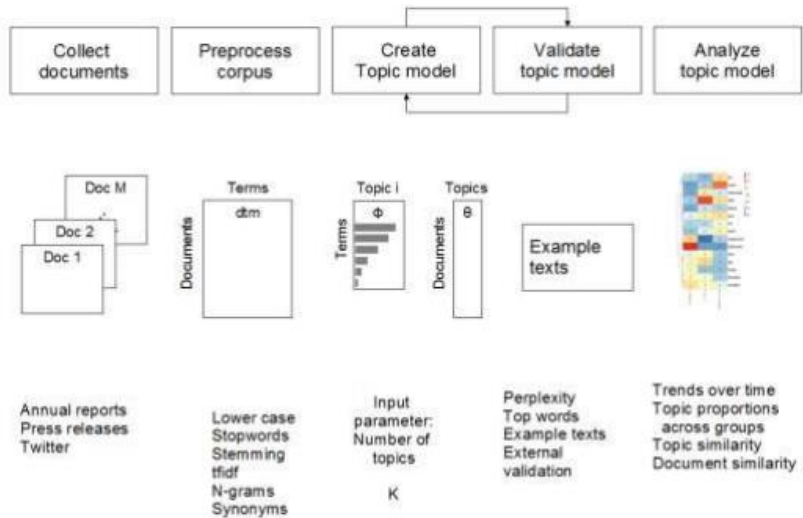
Pada tahap ini dilakukan pemilihan data berdasarkan data yang ada saat ini, yaitu data layanan suara pelanggan. Pemilihan data dapat dimaksudkan sebagai *column selection* atau pemilihan kolom apa saja yang selanjutnya akan dianalisis untuk *topic modelling*. Pemilihan kolom didasarkan pada apa saja yang diinginkan untuk dianalisis sesuai dengan kebutuhan informasi yang ada. Pada penelitian ini, pemilihan kolom didasarkan terhadap kolom deskripsi suara pelanggan, analisa penyebab potensial, tindak lanjut serta penanganan masalah. Dari keempat kolom inilah yang paling menggambarkan topik-topik apa saja yang sering dibahas oleh pelanggan PT Petrokimia Gresik dilihat dari isi konten yang dibicarakan. Baris data dapat diartikan sebagai satu dokumen. Untuk menganalisa dokumen-dokumen baru yang dimasukkan ke dalam database, maka dokumen tersebut akan disesuaikan dengan konten yang dimiliki dimana konten tersebut disesuaikan dengan dokumen yang paling memiliki kemiripan.

3.2.5. Topic Modelling dengan Latent Dirichlet Allocation

Pada tahap ini dilakukan *Topic Modelling* dengan *Latent Dirichlet Allocation* (LDA) terhadap dokumen yang telah dikumpulkan. Sebelum data dapat dianalisis dengan menggunakan LDA, terdapat beberapa langkah-langkah yang perlu dilakukan dalam melakukan *pre-processing* terhadap data.

LDA dilakukan berdasarkan kolom yang telah dipilih pada pemilihan data, untuk kemudian dicari hal-hal apa saja yang sering dibahas pada dokumen. Hasil dari LDA akan merepresentasikan distribusi topik yang paling sesuai pada dokumen suara pelanggan, yang selanjutnya digunakan untuk pembuatan *dashboard* untuk menampilkan informasi apa saja yang sering disampaikan oleh pelanggan, serta bagaimana keterkaitannya dengan performa Bagian Promosi dan Aplikasi Produk yang sedang berjalan, sehingga dapat menjadi pertimbangan untuk menentukan strategi apa yang akan dilakukan oleh pihak promosi ke depannya. Langkah-langkah

dalam melakukan *topic modelling* menggunakan LDA adalah seperti ditampilkan pada Gambar 11.



Gambar 3. 1. Langkah-langkah melakukan LDA[6]

3.2.5.1. Data Preprocessing

Kebanyakan alur *topic modeling* melibatkan langkah-langkah untuk data preprocessing dan pembersihan data. Hal ini akan sangat bergantung pada karakteristik dari data yang akan dianalisis. Pada data Suara Pelanggan di PT. Petrokimia Gresik ini, tahap data preprocessing dan pembersihan data meliputi case folding, tokenization, stemming dan stopword removal. Case folding merupakan tahap untuk membuat kata-kata yang dituliskan pada dokumen menjadi huruf kecil. Hasil dari case folding yang dilakukan pada dokumen suara pelanggan ditampilkan pada Tabel 3.2.

Tabel 3.2. Case Folding

Contoh Data :	Hasil Case Folding :
Penjelasan ke konsumen, PT Petrokimia Gresik tidak bisa melayani pembelian di bawah 30 ton, sehingga konsumen diarahkan ke Distributor PT Petrosida di 031-3981553, atau Layanan Bebas Pulsa 08001401351 atau Layanan SMS 081335060777.	penjelasan ke konsumen, pt petrokimia gresik tidak bisa melayani pembelian di bawah 30 ton, sehingga konsumen diarahkan ke distributor pt petrosida di 031-3981553, atau layanan bebas pulsa 08001401351 atau layanan sms 081335060777.

Setelah melakukan case folding, selanjutnya adalah tahap stemming, yaitu mengubah kata-kata dalam dokumen menjadi kata dasar. Hasil dari stemming ditampilkan dalam Tabel 3.3.

Tabel 3.3. Hasil Stemming pada data

Contoh Data :	Hasil Case Folding :
penjelasan ke konsumen, pt petrokimia gresik tidak bisa melayani pembelian di bawah 30 ton, sehingga konsumen diarahkan ke distributor pt petrosida di 031-3981553, atau layanan bebas pulsa 08001401351 atau layanan sms 081335060777.	jelas ke konsumen, pt petrokimia gresik tidak bisa layan beli di bawah 30 ton, sehingga konsumen arah ke distributor pt petrosida di 031-3981553, atau layan bebas pulsa 08001401351 atau layan sms 081335060777.

Stemming dilanjutkan dengan tahap stopwords removal yaitu menghilangkan kata-kata umum yang tidak memiliki makna

dan tanda baca. Hasil pada stopword removal ditunjukkan pada Tabel 3.4.

Tabel 3.4. Stopword removal pada data

Contoh Data :	Hasil Case Folding :
jelas ke konsumen, pt petrokimia gresik tidak bisa layan beli di bawah 30 ton, sehingga konsumen arah ke distributor pt petrosida di 031-3981553, atau layan bebas pulsa 08001401351 atau layan sms 081335060777.	jelas konsumen pt petrokimia gresik tidak bisa layan beli bawah ton konsumen arah distributor pt petrosida layan bebas pulsa layan sms

Hasil dari stopword removal dilanjutkan dengan melakukan tokenization pada dokumen, yaitu melakukan pemotongan string input berdasarkan tiap kata penyusun dokumen berdasarkan spasi. Hasil dari tokenization ditampilkan pada Tabel 3.5.

Tabel 3.5. Tokenization pada data

jelas	konsumen	pt	petrokimia
gresik	tidak	bisa	layan
beli	bawah	ton	arah
distributor	pt	petrosida	layan
bebas	pulsa	layan	sms

Tahap selanjutnya adalah melakukan pengubahan token dalam dokumen menjadi bentuk corpus. Jika dalam metode text-

mining tahapan ini sering disebut dengan tahapan tfidf. Sedangkan dalam LDA, tahapan ini disebut dengan pembuatan corpus. Hasil dari corpus ditunjukkan pada Tabel 3.6.

Tabel 3.6. Tabel hasil pembentukan korpus

(0, 1)	(1, 2)	(2, 2)	(3, 1)	(4, 1)	(9, 1)
....

3.2.5.2. Pemilihan Model Topik

Tahap ini digunakan untuk memilih jumlah topik yang akan digunakan sebagai vocabulary yang telah didefinisikan pada tahap sebelumnya dengan perplexity yang paling rendah. Melakukan pemilihan model topik dapat dilakukan dengan beberapa cara, diantaranya memperkirakan bahwa kata-kata yang paling banyak muncul pada setiap topik dengan LDAVis *package* dan menilai koherensi dari topik dengan melihat tingkat perplexity. Pemilihan model topik dilakukan dengan menggunakan perplexity. Perplexity merupakan sebuah metode yang digunakan untuk menguji ketepatan atau kesesuaian informasi dari dokumen dengan topik yang dihasilkan. Perplexity mengambil n sampel dari N populasi data untuk diuji, apakah n sampel tersebut memiliki kesesuaian topik dengan kelompok topik dalam N populasi. Selain itu, tahap ini dapat dilakukan dengan melakukan penilaian kesesuaian setiap topik pada dokumen dengan proporsi nilai yang paling tinggi.

3.2.6. Analisis Topik

Interpretasi hasil merupakan sebuah tahap yang dilakukan untuk menganalisis proporsi topik pada korpus. Di dalam sebuah korpus, topik yang paling dominan dan didiskusikan oleh sebagian besar dokumen dihitung dengan menggunakan rata-rata aritmatik pada seluruh dokumen. Hal

ini juga digunakan untuk mengetahui keseragaman distribusi ataupun rata-rata persentase dari cakupan suatu dokumen. Pada tahap ini dilakukan analisis terhadap hasil pengolahan data dengan LDA yang dilakukan untuk membuktikan bahwa distribusi topik yang dihasilkan memiliki kesesuaian pada dokumen suara pelanggan. Sedangkan analisis terhadap hasil LDA dilakukan untuk memperoleh relevansi informasi dengan permasalahan yang diajukan.

3.2.7. Pembuatan *Dashboard*

Pada tahap ini dilakukan pembuatan *dashboard* sehingga dapat menghasilkan kebutuhan informasi yang perlu ditampilkan terhadap Bagian Promosi dan Aplikasi Produk. Tipe *dashboard* yang dibuat adalah tipe *dashboard* operasional dengan *discovery analytics*. *Dashboard* yang dihasilkan akan ditampilkan dan digabungkan pada aplikasi layanan suara pelanggan berbasis web di PT. Petrokimia Gresik.

BAB IV

PERANCANGAN

Pada bab ini akan dijelaskan mengenai perancangan penelitian tugas akhir. Perancangan ini diperlukan sebagai panduan dalam melakukan penelitian tugas akhir, yang dijelaskan sebagai berikut.

4.1. Pengambilan Data

Dalam pelaksanaan analisis dan visualisasi suara pelanggan pada pusat layanan pelanggan PT. Petrokimia Gresik, data merupakan objek utama analisis. Data yang dibutuhkan merupakan data tipe teks hasil konversi Form Suara Pelanggan baik dari media SMS, telepon, dan media cetak.

Pengambilan data dilakukan dari ekstraksi database pusat layanan pelanggan mulai dari Januari 2014 hingga Juni 2016. Atribut pada database layanan suara pelanggan diantaranya adalah No Registrasi, Jenis Masukan, Perihal, Deskripsi, Dari, Akses, Media, Tanggal, Analisa penyebab potensial, Tindak Lanjut, Pencegahan Masalah, Ditindak Lanjuti Oleh, Tanggal, Status. Tabel 4.1. merupakan penjelasan keterangan atribut beserta tipe data dari database pusat layanan pelanggan.

Tabel 4.1. Keterangan Atribut Database

Nama Atribut	Tipe Data	Keterangan
No Registrasi	Int	ID Dokumen Form Suara Pelanggan
Jenis Masukan	varchar(15)	Berisi mengenai jenis masukan pelanggan yaitu berupa pertanyaan,

		informasi atau keluhan
Perihal	varchar(15)	Berisi tentang perihal informasi yang disampaikan yaitu pelayanan, pembelian, produk, harga, promosi atau lainnya.
Deskripsi	text	Berisi tentang suara pelanggan yang telah disampaikan.
Dari	varchar(20)	Berisi tentang identitas yaitu nama pengirim suara pelanggan.
Akses	varchar(15)	Berisi tentang informasi kontak pengirim suara pelanggan.
Media	varchar(15)	Berisi tentang bagaimana pengirim pesan menyampaikan suara pelanggan, yaitu dapat melalui sms, telepon, atau media akses lainnya.

Tanggal	date	Berisi tentang tanggal pengirim suara pelanggan menyampaikan pesan.
Analisa penyebab potensial	text	Berisi tentang analisa sementara permasalahan yang terjadi.
Tindak Lanjut	text	Berisi tentang solusi dari permasalahan yang disampaikan oleh pelanggan.
Pencegahan Masalah	text	Berisi tentang cara penanggulangan permasalahan untuk menghindari permasalahan yang sama terjadi.
Ditindak Lanjuti Oleh	varchar(15)	Unit/departemen yang menyelesaikan permasalahan.
Tanggal	date	Berisi tanggal penyelesaian permasalahan suara pelanggan yang disampaikan.

Status	varchar(10)	Berisi tentang status penanganan permasalahan yang terjadi.
--------	-------------	---

Data yang diekstraksi dari database adalah sejumlah 4400 data. Database yang digunakan Pada Tabel 4.2 merupakan contoh data yang disimpan dalam format xls.

Tabel 4.2. Contoh Data Suara Pelanggan

No	Jenis Masukan	Perihal	Deskripsi	Dari	Akses	Media	Tanggal	Analisa penyebab potensial	Tindak Lanjut	Pencegahan Masalah	Ditindak Lanjut Oleh	Tanggal	Status
1	Keluhan	Produk	SS Jember (Bapak Taufik Muchlisin) mengeluhkan kondisi Petro Cas yang dikirim ke wilayah Kabupaten Jember dalam keadaan rusak, karena produk Petro Cas tidak ditata dengan benar pada saat proses muat.	xxxxxxxxxxxxxxxxxxxxxx	0812-xxxxxxxxxx	SMS	15/03/16	Kurangnya pengawasan pada saat proses pengantongan dan proses muat, sehingga menyebabkan produk Petro Cas dalam kondisi rusak.	Penggantian produk Petro Cas kepada konsumen UD. Rahmat Tani dengan nomor DO 3100015247-B101.	Menambah <i>checker</i> untuk meningkatkan pengawasan pada saat pemuatan produk ke truk pengangkut untuk mengurangi potensi kantong sobek karena handling yang kurang baik. Selain itu, apabila terdapat produk yang cacat untuk tidak dimuat di truk dan diganti dengan produk yang baik.	PPNPJ	02/05/16	Close

2	Pertanyaan	Pembelian	Konsumen menanyakan cara pembelian pupuk Petroganik nonsubsidi.	xxxxxxxxxxxxxxxxxxxxxx	0852-xxxxxxxxxxxxxx	Telepon Bebas Pulsa	01/05/16	Konsumen tertarik untuk menggunakan dan mencoba manfaat produk Petroganik Premium.	Konsumen diarahkan ke Distributor PT Petrosida di 031-3981553, atau Layanan Bebas Pulsa 08001401351 atau Layanan SMS 081335060777.	Menginformasikan ke Distributor PT Petrosida Gresik tentang potensi pasar di wilayah Manado via e-mail.	PLP	01/05/16	Close
3	Pertanyaan	Pembelian	Konsumen menanyakan cara pembelian benih Petro Hybrid Hipa 18.	xxxxxxxxxxxxxx	0856-xxxxxxxxxx	SMS	01/05/16	Konsumen tertarik untuk menggunakan dan mencoba manfaat produk Petro Hybrid Hipa 18.	Konsumen diarahkan untuk menghubungi Departemen Riset (Ibu Eva) di 081223421861.	Menginformasikan ke Departemen Riset tentang potensi pasar di wilayah Kulonprogo.	PLP	01/05/16	Close

4.2.Seleksi Attribut

Setelah melakukan pengambilan data, langkah selanjutnya yang dilakukan adalah melakukan seleksi attribut. Penyeleksian attribut ditujukan untuk memilih attribut apa saja yang akan dianalisa selama proses penelitian berlangsung. Tabel 4.3 akan menampilkan keterangan penyeleksian attribut data.

Tabel 4.3. Hasil Seleksi Atribut Data

Nama Atribut	Seleksi	Keterangan
No Registrasi	X	Tidak digunakan.
Jenis Masukan	X	Tidak digunakan.
Perihal	X	Tidak digunakan.
Deskripsi	V	Digunakan.
Dari	X	Tidak digunakan.
Akses	X	Tidak digunakan.
Media	X	Tidak digunakan.
Tanggal	X	Tidak digunakan.
Analisa penyebab potensial	V	Digunakan.
Tindak Lanjut	V	Digunakan.
Pencegahan Masalah	V	Digunakan.
Ditindak Lanjuti Oleh	X	Tidak digunakan.
Tanggal	X	Tidak digunakan.
Status	X	Tidak digunakan.

Hasil dari tabel tersebut didapatkan bahwa atribut yang digunakan selama penelitian berlangsung adalah atribut deskripsi, analisis penyebab potensial, tindak lanjut serta pencegahan masalah. Keempat atribut ini dipilih karena mengandung teks yang panjang serta berisi tentang konten yang sesuai dengan objek yang akan dianalisa.

Tabel 4.4. merupakan data hasil seleksi atribut yang dipakai untuk penelitian.

Tabel 4.4. Hasil Seleksi Atribut

Deskripsi	Analisa Penyebab Potensial	Tindak Lanjut	Pencegahan Masalah
Menanyakan Harga Pupuk Za, Sp-36, Phonska, Petroganik Dan Superphos?	Konsumen Belum Mengetahui Het Pupuk Bersubsidi.	Penjelasan Ke Konsumen, Harga Eceran Tertinggi Pupuk Bersubsidi Phonska Rp 115.000/Sak, Za Rp 70.000/Sak, Sp-36 Rp 100.000/Sak Dan Petroganik Rp 20.000/Sak. Pembelian Di Kios Resmi Dan Berdasarkan Rdkk, Pembayaran Kontan Serta Diambil Sendiri Di Kios Resmi.	Pengiriman Sticker Het Pupuk Bersubsidi Di Kios-Kios Resmi.
Menanyakan Bagaimana Cara Mendapatkan Pupuk Petroganik Di Riau?	Konsumen Belum Memahami Prosedur Dan Syarat Pembelian Pupuk Bersubsidi.	Penjelasan Ke Konsumen, Untuk Mendapatkan Pupuk Bersubsidi Harus Tergabung Dalam Kelompok Tani Dan Memasukkan Kebutuhan Pupuknya Di Dalam Rdkk (Rencana Difinitif Kebutuhan Kelompok) Yang Disetujui Oleh Dinas Terkait Setempat. Rdkk Digunakan Sebagai	Koordinasi Dengan Base Area Agar Menyisipkan Materi Cara Pembelian Pupuk Bersubsidi Pada Saat Sosialisasi.

		Dasar Untuk Penebusan Pupuk Bersubsidi Di Kios Resmi Terdekat.	
Menanyakan Aplikasi Kapur Pertanian Untuk 1 Hektar Tanaman Cabe?	Konsumen Belum Memahami Cara Aplikasi Kaptan.	Untuk Tanah Pertanian Bersifat Masam, Kapur Pertanian Dosis: 1500-3000 Kg/Hektar, Cara Pemakaian Kapur Pertanian Ditabur Diatas Tanah Pada Saat Pengolahan Tanah Terakhir Atau 2 Minggu Sebelum Tanam.	Koordinasi Dengan Base Area Agar Mensosialisasikan Kembali Cara Pemupukan Berimbang.

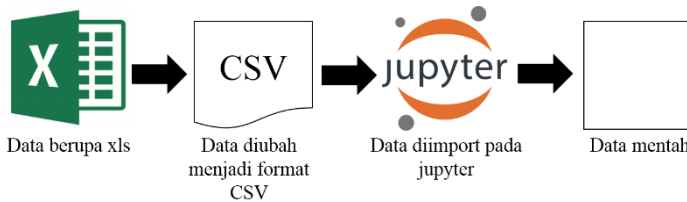
4.3. Metodologi Implementasi Penilaian

Metodologi implementasi penelitian merupakan tahapan yang dilakukan untuk mencapai tujuan penelitian yang disesuaikan dengan komputasi secara otomatis. Komputasi yang dilakukan pada penelitian ini adalah dengan beberapa tahapan, yakni pemrosesan data, pencarian model serta klasifikasi menggunakan python sedangkan visualisasi dashboard menggunakan PHP. Ada 5 tahapan utama dalam melakukan implementasi penelitian yaitu load data, pra-proses data, pemrosesan data, klasifikasi data, dan visualisasi data.

4.3.1. Load Data

Load data merupakan tahap pengambilan data untuk dibaca ke dalam tools sebelum melakukan analisa pada penelitian. Data sebelumnya berformat xls. Kemudian data diubah menjadi

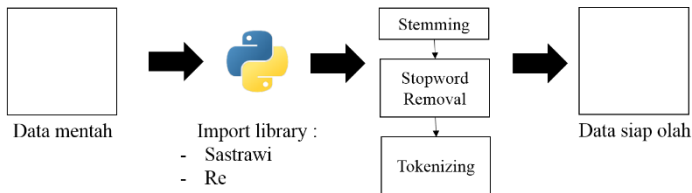
format csv. Pada jupyter notebook, load data dengan format csv dapat dilakukan dengan menggunakan modul csv. Adapun data yang dimuat adalah data keseluruhan suara pelanggan dari bulan Januari 2014 hingga bulan Juni 2016. Gambar 4.1. menampilkan alur dari proses loading data dalam penelitian ini.



Gambar 4.1. Alur Loading Data

4.3.2. Pra-proses Data

Pra-proses data mencakup beberapa langkah utama pengerjaan yakni pengubahan data menjadihuruf kecil, stemming, stopword removal, serta tokenization. Untuk penjelasan secara lebih detail, dijelaskan pada Gambar 4.2.



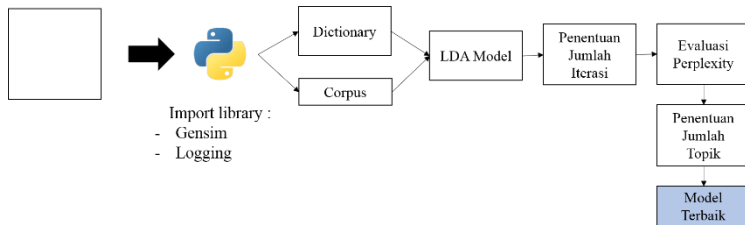
Gambar 4.2. Alur Pra-proses Data

Proses stemming merupakan proses yang dilakukan untuk menghilangkan kata imbuhan serta mengubah kata-kata pada data menjadi kata baku. Pendefinisian stopwords dilakukan dengan mendaftarkan kata-kata dalam bahasa Indonesia, merujuk pada sumber Hasan Alwi yaitu dengan memasukkan kata tugas yang sering digunakan namun tidak memiliki nilai informasi. Sedangkan tokenization dilakukan untuk menghilangkan penulisan huruf besar menjadi huruf kecil serta

menghilangkan adanya karakter yang tidak diperlukan. Pada python, kedua tahapan ini dapat dilakukan dengan modul `re` dan `defaultdict`. Kemudian, dari hasil tokenization, data diterjemahkan untuk menjadi corpus serta disimpan untuk penggunaan pemrosesan data selanjutnya.

4.3.3. Proses Data

Pada tahapan proses data, langkah utama yang dilakukan adalah melakukan pencarian model dengan menggunakan modul LDA. Gambar 4.3. menjelaskan tentang bagaimana alur pemrosesan data dijalankan.

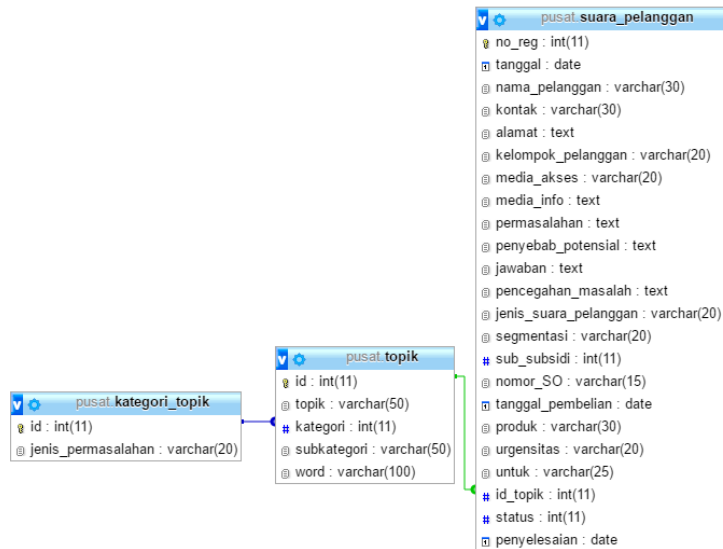


Gambar 4.3. Alur Pemrosesan Data

Hasil dari pencarian model akan digunakan untuk mendapatkan topik apa saja yang muncul dari analisa pada dokumen. Untuk melakukan evaluasi terhadap model, menggunakan rata-rata perplexity atau dengan mengukur kebingungan suatu dokumen jika dikelompokkan dalam topik tertentu. Evaluasi perplexity terdapat dalam python dengan modul logging. Hasil perplexity yang lebih kecil serta mulai stabil atau tidak berubah-ubah, digunakan sebagai pilihan model yang akan digunakan.

4.3.4. Analisis Topik

Analisis topik digunakan untuk mengidentifikasi hasil topik yang diterjemahkan dari proses LDA terhadap jumlah topik yang dipilih. Gambar 4.4. menampilkan alur dari analisis topik dalam penelitian ini.

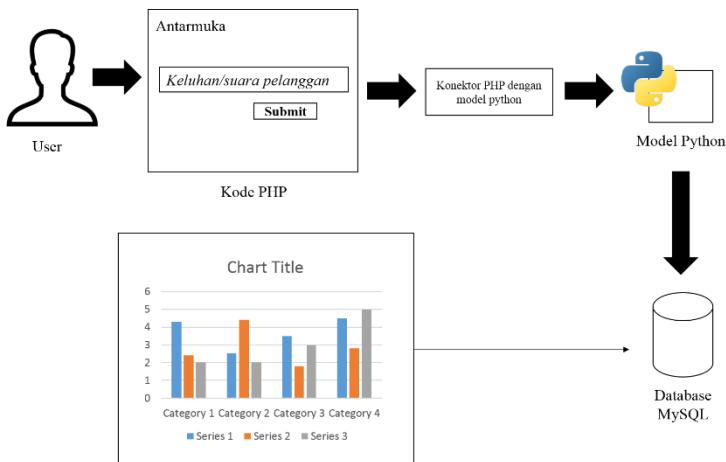


Gambar 4.5. Perancangan Database Aplikasi Pusat Layanan Pelanggan

Dalam perancangan database terdapat tabel suara_pelanggan sebagai penyimpanan transaksi yang dilakukan untuk form suara pelanggan. Pada tabel suara pelanggan, terdapat atribut diantaranya no_reg, tanggal, suara_pelanggan, kontak, alamat, kelompok_pelanggan, media_akses, media_info, permasalahan, penyebab_potensial, jawaban, pencegahan_masalah, id_jenis_suara_pelanggan, segmentasi, sub_subsidi, nomor_SO, tanggal_pembelian, produk, urgensi, untuk, serta id_topik. Kolom id_topik merupakan kolom yang digunakan sebagai pengembalian hasil pemrosesan data menggunakan model python. Sedangkan tabel topik merupakan tabel yang berisi struktur topik berupa id_topik, topik, kategori_topik, serta sub_kategori_topik. Hasil dari penghitungan jumlah id_topik dan kategori_topik yang kemudian digunakan untuk visualisasi dashboard pada Aplikasi Pusat Layanan Pelanggan.

4.3.6.2. Menghubungkan PHP dan Python

Pada tahapan ini dilakukan penghubungan data dari input melalui kode PHP dengan model python untuk kemudian digunakan sebagai proses klasifikasi data. Penjelasan secara detail dari rancangan konektivitas model python dengan kode PHP ditampilkan pada Gambar 4.6. User dapat menginputkan data melalui antarmuka aplikasi, kemudian ketika user menekan tombol submit, merupakan trigger untuk menjalankan model python. Hasil dari model python kemudian disimpan kedalam database. Hasil dari database kemudian divisualisasikan menggunakan kode PHP dengan bar chart.



Gambar 4.6. Alur menghubungkan antarmuka dengan model python

4.3.6.3. Desain Antarmuka Visualisasi Dashboard

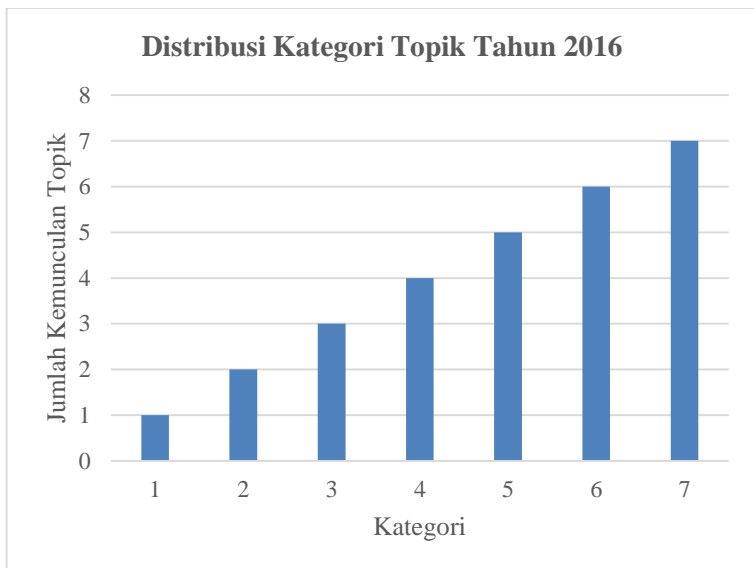
Setelah mendapatkan model yang tepat dari hasil klasifikasi, selanjutnya dilakukan penghitungan berdasarkan kategori data yang telah diklasifikasikan ke dalam topik sesuai dengan waktu yang ditentukan. Hasil penghitungan jumlah dokumen ke dalam topik tersebut kemudian divisualisasikan ke dalam dashboard dengan menggunakan bahasa pemrograman PHP. Rancangan *dashboard* yang akan ditampilkan dari analisa

topic modelling, dibedakan berdasarkan konten informasi yang akan ditampilkan. Pada Gambar 4.7. menjelaskan tentang dashboard jumlah suara pelanggan yang masuk, jumlah suara pelanggan tertangani dan jumlah pelanggan aktif.

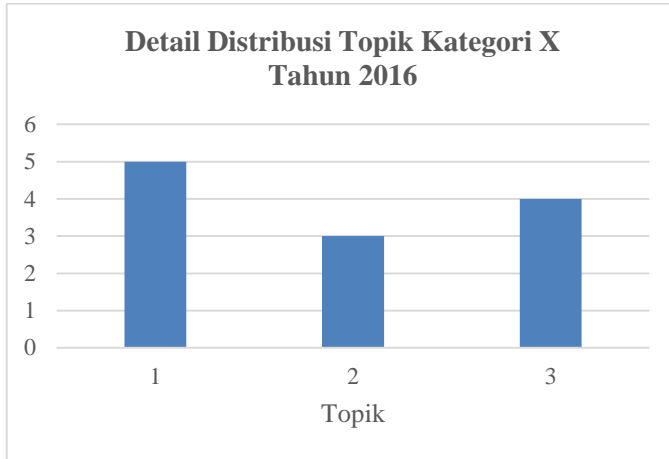
Jumlah Suara Pelanggan Masuk	Jumlah Suara Pelanggan Tertangani	Jumlah Pelanggan Aktif
120	120	50

Gambar 4.7. Jumlah suara pelanggan dan jumlah pelanggan

Gambar 4.8. menampilkan rancangan grafik distribusi topik tahunan. Grafik diambil dari jumlah kemunculan topik pada dokumen yang dikelompokkan berdasarkan kategori. Data diambil dari data sesuai dengan tahun saat ini.



Gambar 4.8. Distribusi Topik Tahunan



Gambar 4.9. Detail distribusi topik berdasarkan kategori

Gambar 4.9. menampilkan rancangan detail grafik distribusi topik berdasarkan kategori. Grafik diambil dari jumlah kemunculan topik pada dokumen yang dikelompokkan berdasarkan topik pada setiap kategori. Data diambil dari data sesuai dengan tahun saat ini.

Tabel 4.5. menampilkan rancangan tampilan tabel untuk menunjukkan kabar terbaru suara pelanggan. Tabel akan menunjukkan nama pelanggan, deskripsi, kategori topik serta tanggal suara pelanggan tersebut diinputkan ke dalam aplikasi Pusat Layanan Pelanggan. Tabel menampilkan setiap kabar terbaru dengan jumlah 10 teratas.

Tabel 4.5. Kabar Terbaru Suara Pelanggan

No	Nama Pelanggan	Deskripsi	Kategori Topik	Tanggal
1	xxxx	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXX	1	10/12/2016
2	yyyy	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXX	2	10/12/2016
.
.
10	aaaaa	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXX	1	9/12/2016

Tabel 4.6. menampilkan rancangan tampilan tabel untuk menunjukkan kabar terbaru suara pelanggan berdasarkan kategori yang dipilih. Tabel akan menunjukkan nama pelanggan, deskripsi, topik serta tanggal suara pelanggan tersebut diinputkan ke dalam aplikasi Pusat Layanan Pelanggan. Jumlah kabar terbaru yang ditampilkan pada setiap tabel adalah berjumlah 10 teratas. Topik menyesuaikan dengan kategori topik yang telah dipilih oleh pengguna.

Tabel 4.6. Kabar terbaru suara pelanggan kategori X

No	Nama Pelanggan	Deskripsi	Topik	Tanggal
1	xxzszzz	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXX	a	10/12/2016
2	zdsdkd	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXX	b	10/12/2016
.
.
10	cbncbjbjh	XXXXXXXXXXXXXXXX XXXXXXXXXXXXXX	a	9/12/2016

Halaman sengaja dikosongkan

BAB V

IMPLEMENTASI

Bab ini menjelaskan hasil dari implementasi perancangan studi kasus atau hasil dari proses pelaksanaan penelitian. Hasil yang akan dijabarkan adalah hasil eksperimen terhadap data yang digunakan sebagai acuan penelitian. Selain itu, akan dijelaskan juga mengenai hambatan dan rintangan dalam proses pelaksanaan penelitian.

5.1.Perangkat Penelitian

Dalam pelaksanaan identifikasi topik dan visualisasi data suara pelanggan PT. Petrokimia Gresik, dibutuhkan perangkat-perangkat untuk menunjang keberlangsungan penelitian. Adapun perangkat-perangkat yang dibutuhkan berupa perangkat keras dan perangkat lunak, dengan spesifikasi sebagai berikut :

Hardware

- CPU : Intel® Core™ i3-3110M CPU @ 2.40GHz
2.40 GHz
- Monitor :14 inch 1366x768
- Keyboard

Software

- Operating system : Windows 8.1 64-bit
- Bahasa pemrograman : Python 3.5 64-bit, PHP
- Tools : Anaconda 4.2.0
Ms. Excel 2013
Notepad++
MySQL
XAMPP

5.2.Memuat Data

Dalam melakukan loading data atau memuat data secara keseluruhan, pada penelitian ini menggunakan jupyter notebook. Data yang dimuat memiliki format .csv. Untuk dapat melakukan pembacaan format CSV pada jupyter,

menggunakan modul CSV. Data diambil dari beberapa file yaitu data mentah dan data hasil stemming. Kode 5.1. merupakan kode yang digunakan dalam proses memuat data mentah. Proses memuat data ini selanjutnya digunakan sebagai input atau bahan untuk melakukan pra-proses data.

```
import csv
x = open('data_mentah.csv')
csv_x = csv.reader(x)
for line in doc:
    documents.append(line)
```

Kode 5.1. Loading Data

5.3.Pengerjaan Pra-Proses Data

Pra-pemrosesan data dalam analisis suara pelanggan menggunakan pemodelan topik merupakan salah satu tahapan yang penting dalam penelitian. Pra-pemrosesan data meliputi beberapa tahapan diantaranya pengubahan data suara pelanggan menjadi huruf kecil, stemming, stopword removal, tokenization.

5.3.1. Stemming

Stemming merupakan tahapan untuk mengubah data menjadi kata dasar. Dalam penelitian ini, stemming dilakukan dengan menggunakan library Sastrawi dengan fungsi stem. Stemming perlu dilakukan karena dalam penelitian, data yang digunakan sebagian besar merupakan kata formal. Untuk melakukan stemming pada data, Kode 5.2. merupakan kode yang digunakan. Data yang di stem merupakan data setiap baris yang terdapat dalam dokumen yang berjumlah 4400 baris. Setelah data berhasil di stem, data kemudian disimpan ke dalam suatu file csv baru untuk kemudian digunakan pada tahapan selanjutnya. Dari hasil stemming kemudian dilakukan pengecekan kembali secara manual jika terdapat beberapa kesalahan pengubahan kata dasar.

```

from Sastrawi.Stemmer.StemmerFactory import
StemmerFactory

factory = StemmerFactory()
stemmer = factory.create_stemmer()

array=[]
writer =
csv.writer(open("data_siap.csv","wb"),delimiter=" ")

for row in csv_x:
    output = stemmer.stem(row[0])
    array.append(output.split())

for data in array:
    writer.writerow(data)

x.close()

```

Kode 5.2. Stemming pada data suara pelanggan

5.3.2. Pendefinisian Stopword

Stopword merupakan tahapan yang digunakan untuk menghilangkan kata-kata yang tidak bermakna dalam data. Stopword dilakukan dengan menggunakan kata tugas berdasarkan Tata Bahasa Baku Bahasa Indonesia dari Hasan Alwi. Selain kata tugas, dalam stopwords juga ditambahkan beberapa kata-kata yang tidak memiliki makna dari suara pelanggan, seperti http, .com, bu, bapak, ibu. Kode yang digunakan untuk melakukan stopwords removal ditampilkan pada Kode 5.3. Data setiap baris yang telah berhasil melalui stopwords removal kemudian disimpan dalam sebuah list.

```

stoplist = {"akan", "bersama", "bagaikan", "melalui",
"mengenai", "dari pada", "di", "ke", "dari", "hingga",
"sampai", "bagi", "untuk", "guna", "buat", "dengan",
"sambil", "beserta", "bersama", "karena", "sebab",
"lantaran", "oleh", "tentang", "mengenai", "dari", "dan",
"serta", "atau", "tetapi", "melainkan", "padahal",
"sedangkan", "sejak", "semenjak", "sedari", "sewaktu",
"ketika", "tatkala", "selama", "demi", "serta", "hingga",
"sampai", "setelah", "sesudah", "sebelum", "selesai",
"seusai", "sehabis", "jika", "kalau", "jikalau", "bila",
"manakala", "andaikan", "umpamanya", "biarpun",
"sekalipun", "seakan-akan", "seperti", "sebagai",
"sebab", "karena", "oleh sebab", "sehingga", "sampai",
"dengan", "tanpa", "bahwa", "yang", "sama dengan",
"lebih dari", "bah", "cih", "cis", "ih", "idih", "brengsek",
"sialan", "buset", "keparat", "aduhai", "amboi", "asyik",
"syukur", "alhamdulillah", "insya allah", "aduh", "aih",
"ai", "lo", "duilah", "eh", "oh", "ah", "astaga",
"astagfirullah", "masyaallah", "ayo", "mari", "hai", "be",
"eh", "halo", "nah", "sang", "sri", "hang", "dang", "si",
"bagaimana", "mengapa", "apa", "siapa", "dimana",
"kapan", "menjadi", "menanyakan", "dalam", "http",
"ini", "itu", "com", "ingin", "ha", "lebih", "rp",
"merupakan", "adalah", "pada", "bapak", "ibu", "bpk",
"bu"}

```

```

texts = [[word for word in document.lower().split() if
word not in stoplist]for document in
documents]

```

Kode 5.3. Stopword Removal

5.3.3. Tokenization

Tokenization merupakan langkah-langkah dalam menghilangkan karakter dan angka serta memisahkan kalimat menjadi kata-kata. Tokenization merupakan salah satu tahapan dimana data dipisah berdasarkan spasi untuk dijadikan kata dalam korpus. Data di filter berdasarkan jumlah frekuensi yang terdapat pada setiap kata.

```
import re
tokens = []
for word in texts:
    clean = []
    for i in word:
        for y in re.split('[^a-zA-Z0-9]+',i):
            if ((len(y)>1) and re.match('[a-zA-Z]',y)):
                clean.append(y)
    tokens.append(clean)

from collections import defaultdict
frequency = defaultdict(int)
for text in tokens:
    for word in text:
        frequency[word] += 1

tokens = [[word for word in text if frequency[word] >
1] for text in tokens]
```

Kode 5.4. Tokenization pada suara pelanggan

Kode 5.4. menampilkan bagaimana proses tokenization dilakukan. Tokenization pada penelitian ini menggunakan modul re dan defaultdict yang terdapat dalam gensim. Modul re digunakan untuk memisahkan dan menghilangkan angka ataupun karakter yang terdapat dalam dokumen yang telah disimpan ke dalam array texts pada tahap sebelumnya.

Dari hasil pra proses data berupa menghilangkan karakter dan angka, selanjutnya melakukan tahapan menghitung jumlah kata yang terdapat dalam array text. Kata dihitung kemunculannya dalam setiap dokumen suara pelanggan. Jika ditemukan jumlah frekuensi kata yang ada melebihi 1 kali, maka kata tersebut selanjutnya dimasukkan ke dalam array tokens. Sedangkan jika kata tersebut tidak melebihi 1 kali frekuensi kemunculannya, maka kata tersebut tidak turut diinputkan ke dalam array.

5.4.Pengerjaan Proses Data

Tahap pemrosesan data dimulai dengan melakukan penyimpanan dictionary dari setiap kata yang terdapat dalam dokumen yang telah di lakukan pra prosesing. Kode untuk melakukan pembuatan dictionary ditunjukkan pada Kode 5.5.

```
dictionary = corpora.Dictionary(tokens)
dictionary.save('dictionary.dict')
```

Kode 5.5. Pembuatan Dictionary

Fungsi dictionary() adalah untuk memberikan nilai unik seperti index berupa integer pada setiap kata untuk mempermudah proses pada tahapan selanjutnya. Hasil dari dictionary yang telah diproses kemudian disimpan ke dalam file bernama dictionary.dict untuk dapat digunakan pada tahapan-tahapan selanjutnya.

Untuk dapat mengetahui isi dan melakukan pengecekan pada dictionary yang telah terbentuk dapat dilakukan fungsi print seperti pada Kode 5.6.

```
print(dictionary)
```

Kode 5.6. Print dictionary

Sedangkan untuk mengetahui jumlah setiap kata yang muncul dalam dokumen yang telah disimpan pada dictionary adalah dengan menggunakan Kode 5.7.

```
print(dictionary.token2id)
```

Kode 5.7. Print jumlah kata yang terdapat dalam dictionary

Dari hasil pembuatan dictionary, langkah yang harus dilakukan selanjutnya adalah membuat korpus atau kumpulan vektor kata dari dokumen. Untuk membuat kata menjadi korpus, menggunakan fungsi doc2bow.

```
corpus = [dictionary.doc2bow(text) for text in tokens]
corpora.MmCorpus.serialize('corpora.mm', corpus)
```

Kode 5.8. Membuat Korpus

Hasil korpus perlu disimpan untuk pembuatan dan penggunaan model secara lebih lanjut. Untuk dapat melakukan pembuatan corpus, maka kode yang perlu dituliskan adalah seperti pada Kode 5.8.

Korpus kemudian di proses untuk membuat model LDA yang paling sesuai. Model yang memiliki kesesuaian lebih besar, dilihat dari hasil perplexity yang paling kecil.

5.4.1. Penentuan Iterasi

Penentuan iterasi merupakan salah satu tahapan yang krusial untuk menentukan model yang tepat untuk digunakan. Iterasi merupakan perulangan yang digunakan dalam training data menggunakan LDA. Istilah lain iterasi dalam LDA adalah passes. Penentuan jumlah passes ini adalah berdasarkan eksperimen yang dilakukan sebanyak 5 kali dengan jumlah topik sebanyak 5, 10 dan 15. Passes yang diamati adalah sejumlah 30 passes. Dari 30 passes yang diamati kemudian dilihat titik kestabilan dari model Untuk dapat melakukan eksperimen tersebut, maka kode yang perlu dijalankan terdapat dalam Kode 5.9.

```

import logging
info = logging.basicConfig(format='%(asctime)s :
%(levelname)s : %(message)s', level=logging.INFO)
from gensim import corpora, models, similarities
from gensim.models import LdaModel
for i in range (3):
    lda_a = LdaModel(corpus,
                      num_topics=5, passes=30, alpha='auto',
                      eval_every=5)

```

Kode 5.9. Percobaan 30 iterasi

Dari hasil perplexity yang muncul, kemudian di dokumentasikan untuk dapat melihat perbandingan setiap nilai yang telah di-*generate* oleh model.

5.4.2. Penentuan Jumlah Topik

Setelah mengetahui jumlah iterasi yang tepat sehingga membuat model menjadi lebih stabil, tahapan selanjutnya yang perlu dilakukan adalah menentukan jumlah topik yang tepat. Topik diperoleh berdasarkan beberapa pertimbangan, diantaranya adalah pertimbangan kegunaan atau *usefulness* dan kualitas model. Berdasarkan pertimbangan kegunaan, terlalu banyak topik tidak membuat semakin singkat proses, justru dapat membuat lebih rumit. Hal tersebut kontradiktif dengan jumlah topik yang terlalu sedikit. Membuat jumlah topik terlalu sedikit bisa saja tidak dapat mencakup topik-topik yang diinginkan.

Sedangkan penentuan jumlah topik berdasarkan kualitas model adalah dengan mengukurnya menggunakan perplexity. Perplexity merupakan suatu kebingungan menggolongkan atau mengelompokkan suatu dokumen baru ke dalam suatu topik tertentu yang telah dimiliki sebelumnya.

Penentuan jumlah topik ini dilakukan pada rentang 3-50 topik. Rentang nilai tersebut dianggap sesuai dengan representasi pertimbangan yang telah dibahas sebelumnya. Percobaan pada setiap jumlah topik dilakukan sebanyak 10 kali. Untuk menjalankan proses tersebut, maka kode yang perlu dijalankan ditampilkan dalam Kode 5.10. Kode tersebut dijalankan sebanyak 10 kali.

```
lda_3 = Idamodel.LdaModel(corpus,
                           id2word=dictionary, num_topics=3,
                           passes=20, alpha='auto', eval_every=5)
lda_3.save('3.model')
```

Kode 5.10. LDA Model dengan jumlah topik 3

5.5. Analisis Topik

Analisis topik merupakan tahapan yang dilakukan berdasarkan luaran dari LDA Model yang telah dipilih. Analisis topik dilakukan dengan mengeluarkan semua kemungkinan topik serta distribusi kata-kata dalam topik tersebut. Untuk menunjukkan hasil topik, dapat menggunakan Kode 5.11. Fungsi `print_topics(-1)` merupakan sebuah fungsi yang digunakan untuk mengurutkan probabilitas distribusi kata yang terdapat dalam topik.

```
lda_35.print_topics(-1)
```

Kode 5.11. Print hasil topik dan distribusi kata

Topik yang dihasilkan kemudian dianalisa serta disesuaikan dengan kebijakan yang terdapat dalam perusahaan.

5.6. Klasifikasi Data

Topik yang telah didefinisikan pada tahapan analisis topik digunakan sebagai input untuk melakukan klasifikasi. Klasifikasi merupakan pelabelan dokumen dengan topik-topik yang telah ditentukan. Klasifikasi dilakukan pada keseluruhan

dokumen. Untuk melakukan klasifikasi terhadap data, Kode 5.12 merupakan kode yang digunakan.

```
topic_mixture_list_10 = [dict(lda_10[x]) for x in corpus]
for topics in topic_mixture_list_10:
    print(topics)
    print (max(topics, key=topics.get))
dict_10 = {i: max(item, key=item.get) for i, item in
enumerate(topic_mixture_list_10)}

from collections import defaultdict
v_10 = defaultdict(list)
for key, value in sorted(dict_10.items()):
    v_10[value].append(key)

for topic_index in v_10.keys():
    print(topic_index)
```

Kode 5.12. Klasifikasi topik

5.5.Konektor PHP dengan Python

Sebelum melakukan visualisasi data, tahapan yang juga perlu dilakukan adalah mengkoneksikan model python yang telah dibuat dengan PHP. Hal ini bertujuan untuk proses kedepannya, jika terdapat masukan pada form suara pelanggan, maka dokumen akan secara otomatis diklasifikasikan pada topik-topik yang telah didefinisikan. Kode pada PHP digunakan sebagai trigger untuk memproses model python serta menghasilkan klasifikasi topik yang akan dikembalikan ke dalam database. Untuk melakukan koneksi dari input pada PHP dengan model python yang harus dijalankan, dapat digunakan dengan menggunakan Kode 5.13.

```
$result = exec('C:\Users\A46CB\Anaconda3\python
model.py "' . $_POST['permasalahan'].'", $err);
```

Kode 5.13. Penghubung model python dengan kode PHP

5.6.Pengerjaan Visualisasi Data

Visualisasi data dalam penelitian ini merupakan dashboard yang dihasilkan dari suara pelanggan yang masuk melalui aplikasi layanan suara pelanggan. Dashboard dibuat untuk menunjukkan trend topik yang sering dibahas oleh pelanggan. Visualisasi data ke dalam dashboard dibuat dengan menggunakan bahasa pemrograman PHP. Visualisasi berikut ditampilkan dengan menggunakan grafik dan tabel. Grafik ditampilkan dengan menggunakan bar chart dengan menggunakan library highchart.

Pada Gambar 5. 1. dapat dijelaskan bahwa pada antarmuka aplikasi, terdapat form suara pelanggan. Form tersebut digunakan sebagai bahan input data yang masuk ke dalam database serta dijalankan oleh model python.

Pada Gambar 5.2. dapat dijelaskan bahwa pada antarmuka aplikasi, bagian kotak atas menampilkan fungsi untuk dapat melihat jumlah suara pelanggan yang terkumpul, jumlah suara pelanggan yang tertangani, jumlah pelanggan aktif. Sedangkan grafik dibawah kotak merupakan fungsi yang digunakan untuk melihat distribusi topik tahunan yang dikelompokkan untuk setiap kategori.

Pada Gambar 5.3. menjelaskan antarmuka aplikasi yang dapat menampilkan secara lebih detail dari kategori topik tahunan yang telah dipilih sebelumnya.

Pada Gambar 5.4. menjelaskan antarmuka aplikasi berupa fungsi untuk melihat kabar terbaru seluruh kategori. Kabar terbaru ditampilkan dalam tabel dengan jumlah 20 baris. Dalam kabar terbaru terdapat fungsi view detail untuk melihat lebih detail mengenai suara pelanggan yang masuk.

Pada Gambar 5.5. menjelaskan antarmuka aplikasi berupa tampilan detail dari kabar terbaru yang dipilih. Terdapat menu untuk mengedit topik yang tidak sesuai.

Pada Gambar 5. 6. menjelaskan antarmuka aplikasi yang digunakan untuk mengedit topik yang dihasilkan oleh model jika terdapat ketidakcocokan sesuai dengan yang diharapkan.

Pada Gambar 5.8. menjelaskan antarmuka aplikasi berupa fungsi untuk melihat distribusi topik tiap kategori bulanan yang dipilih sebelumnya.

Pada Gambar 5. 7. menjelaskan antarmuka aplikasi berupa konfirmasi untuk menentukan topik dari suatu dokumen. Jika benar maka topik adalah sesuai dari hasil model sedangkan jika salah, pengguna aplikasi dapat menentukan topik yang sesuai dengan dokumen.

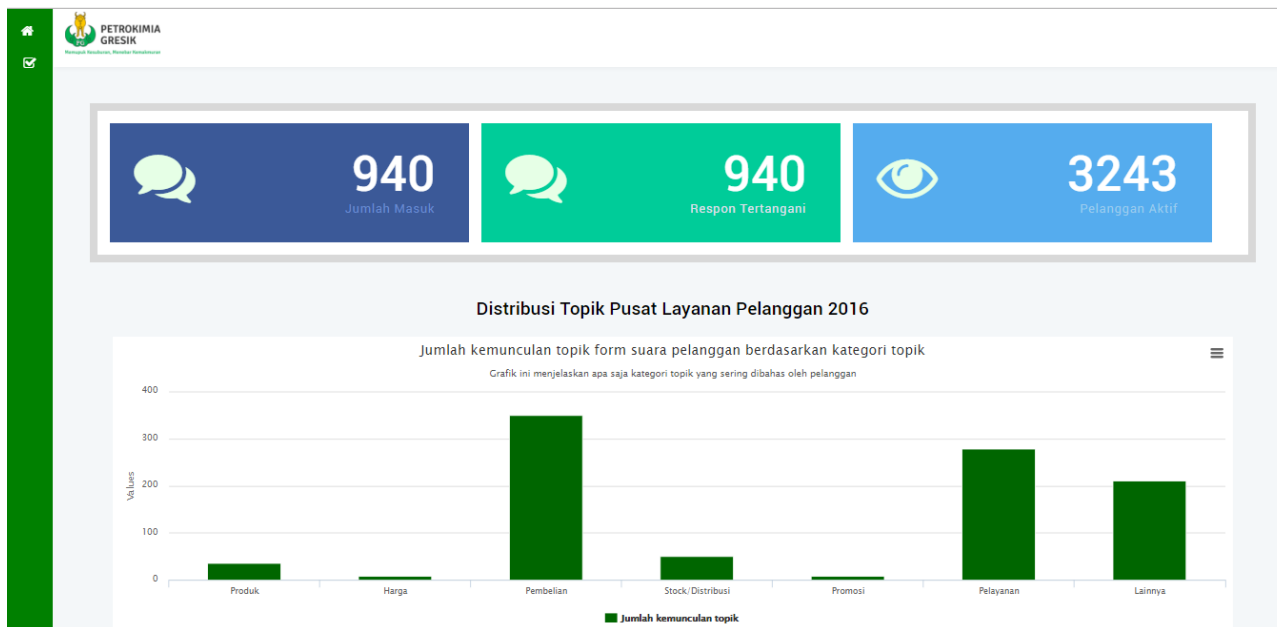
Pada Gambar 5.9. menjelaskan antarmuka aplikasi berupa fungsi untuk melihat detail distribusi topik berdasarkan kategori bulanan yang dipilih sebelumnya.

Pada Gambar 5.10. menjelaskan antarmuka aplikasi berupa fungsi untuk melihat total distribusi topik berdasarkan kategori pada setiap bulan.

Form Suara Pelanggan

Tujuan	<input type="text" value="Departemen yang dituju"/>
Level Suara Pelanggan	<input type="radio"/> Urgent <input type="radio"/> Normal
Jenis Suara Pelanggan	<input type="radio"/> Informasi <input type="radio"/> Pertanyaan <input type="radio"/> Keluhan
Nama/Instansi	<input type="text" value="Nama Pelanggan/Instansi"/>
Alamat	<input type="text" value="Alamat : Jalan xxx, Gresik, Jawa Timur"/>
Kontak	<input type="text" value="Nomor Telepon/Email/Kontak"/>
Segmen Pelanggan	<input type="text" value="Subsidi"/>
Subsidi	<input type="checkbox"/> RDKK <input type="checkbox"/> Kelompok Tani <input type="checkbox"/> Alokasi Pupuk Bersubsidi
Nomor SO	<input type="text" value="Nomor Sales Order"/>
Tanggal Pembelian	<input type="text" value="Format : yyyy-mm-dd"/>
Nama Produk	<input type="text" value="Contoh : Kapur Pertanian, Urea"/>
Kelompok Pelanggan	<input type="radio"/> Corporate <input type="radio"/> Retail

Gambar 5. 1. Form Suara Pelanggan



Gambar 5.2. Halaman Utama Aplikasi

Detail Distribusi Kategori Topik Pembelian Pusat Layanan Pelanggan 2016



Gambar 5.3. Detail Distribusi per Kategori Topik

Kabar Terbaru

Berikut merupakan daftar suara pelanggan yang masuk disertai dengan kategori topik:

NO	WAKTU	NAMA PELANGGAN	DESKRIPSI	KATEGORI TOPIK	DETAIL
1	2017-01-11		Konsumen menanyakan cara pembelian pupuk non subsidi	Stock/Distribusi	View Detail
2	2017-01-11		Konsumen menanyakan cara pembelian pupuk non subsidi	Pelayanan	View Detail
3	2017-01-11		Konsumen menanyakan cara pembelian pupuk petroganik nonsubsidi	Lainnya	View Detail
4	2017-01-10	Sistem Informasi	Konsumen menanyakan cara pembelian pupuk bersubsidi	Pembelian	View Detail
5	2016-06-30	Eno	Konsumen menanyakan cara pembelian benih Petro Hybrid Hipa 18.	Pembelian	View Detail
6	2016-06-29	Hermina	Konsumen menanyakan cara pembelian Kaptan sejumlah 22 ton.	Pembelian	View Detail
7	2016-06-29	Husein	Konsumen menanyakan cara pembelian benih Petro Hybrid Hipa 18.	Pembelian	View Detail
8	2016-06-29	Ahmad Farid	Konsumen menanyakan tentang recruitment karyawan PT Petrokimia Gresik.	Pelayanan	View Detail
9	2016-06-29	Kompas.com	Wajah Gembira Para Tukang Becak di Gresik-Raut wajah gembira terpancar dari para tukang becak di Kabupaten Gresik, Jawa Timur. Mereka mendapat bantuan dari PT Petrokimia Gresik menjelang Hari Raya Idul Fitri 1437 Hijriah. Dalam kegiatan yang berlangsung di Gedung Olahraga Tri Dharma, Gresik, Selasa (28/6/2016) ini. Sebanyak 1.300 tukang becak yang biasa mangkal di wilayah Gresik mendapat bantuan dari PT PG. Santunan tersebut berupa uang tunai sebesar Rp 100.000, paket konsumsi untuk buka puasa, dan selebar sarung kepada setiap tukang becak yang hadir. Diharapkan bantuan ini dapat meringankan beban abang tukang becak dan keluarganya dalam menyambut hari raya Idul Fitri 1437 Hijriah.	Lainnya	View Detail

Gambar 5.4. Kabar Terbaru Layanan Suara Pelanggan

Detail Topik Kabar Terbaru

TANGGAL PELAPORAN	2016-06-29
NAMA PELANGGAN	Kompas.com
DESKRIPSI	Wajah Gembira Para Tukang Becak di Gresik-Raut wajah gembira terpancar dari para tukang becak di Kabupaten Gresik, Jawa Timur. Mereka mendapat bantuan dari PT Petrokimia Gresik menjelang Hari Raya Idul Fitri 1437 Hijriah. Dalam kegiatan yang berlangsung di Gedung Olahraga Tri Dharma, Gresik, Selasa (28/6/2016) ini. Sebanyak 1.300 tukang becak yang biasa mangkal di wilayah Gresik mendapat bantuan dari PT PG. Santunan tersebut berupa uang tunai sebesar Rp 100.000, paket konsumsi untuk buka puasa, dan selempar sarung kepada setiap tukang becak yang hadir. Diharapkan bantuan ini dapat meringankan beban abang tukang becak dan keluarganya dalam menyambut hari raya Idul Fitri 1437 Hijriah.
ANALISIS PENYEBAB POTENSIAL	Menjelang lebaran, PT Petrokimia Gresik berbagi kebahagiaan dengan abang becak yang berada di sekitar perusahaan.
TINDAK LANJUT	Dikliping.
PENCEGAHAN MASALAH	Tidak perlu pencegahan masalah, karena merupakan berita positif tentang Petrokimia Gresik.
KATEGORI TOPIK	Lainnya
TOPIK	Kabar Positif Perusahaan
Edit	

Gambar 5. 5. Detail Kabar Terbaru Layanan Suara Pelanggan

Detail Topik Kabar Terbaru	
TANGGAL PELAPORAN	2017-01-11
NAMA PELANGGAN	
DESKRIPSI	Konsumen menanyakan cara pembelian pupuk non subsidi
ANALISIS PENYEBAB POTENSIAL	Konsumen tidak mengerti cara pembelian pupuk nonsubsidi
TINDAK LANJUT	Penjelasan tentang cara pembelian pupuk nonsubsidi
PENCEGAHAN MASALAH	Penjelasan ke konsumen
TOPIK	<input type="text" value="Alamat Distributor"/>
<input type="button" value="Simpan"/>	

Gambar 5. 6. Edit Topik Kabar Terbaru

Entri suara pelanggan ini akan dimasukkan ke topik : Harga Pupuk ✕

☒ Benar
☐ Salah, pilih topik :

Gambar 5. 7. Edit Topik Kabar Terbaru Setelah Pengisian Form Suara Pelanggan

Distribusi Topik Pusat Layanan Pelanggan 2016

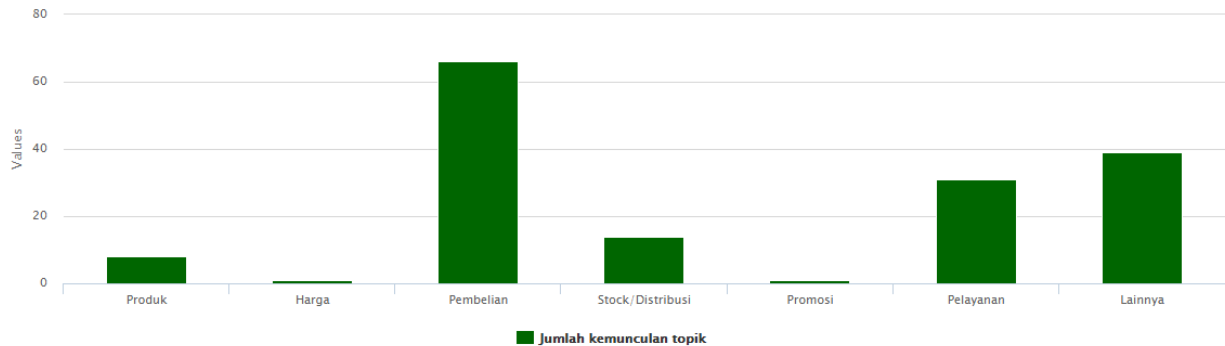
Bulan

Januari

Pilih

Jumlah kemunculan topik form suara pelanggan berdasarkan kategori topik

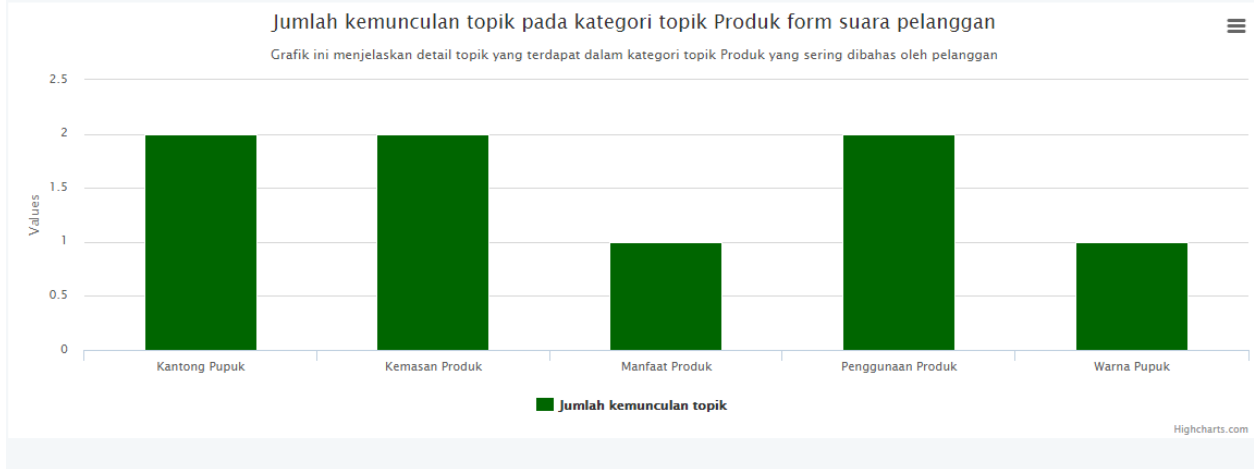
Grafik ini menjelaskan apa saja kategori topik yang sering dibahas oleh pelanggan



Highcharts.com

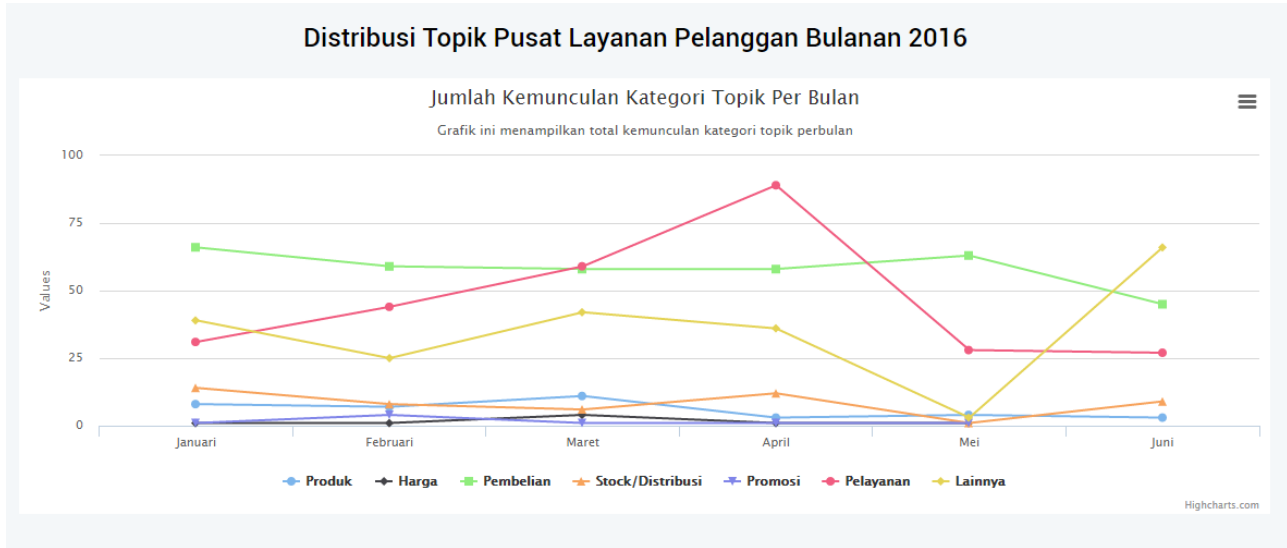
Gambar 5.8. Distribusi Topik Berdasarkan Kategori per Bulan

Detail Distribusi Topik Produk Pusat Layanan Pelanggan Januari 2016



Gambar 5.9. Distribusi Detail Topik per Kategori dan Bulan

Distribusi Topik Pusat Layanan Pelanggan Bulanan 2016



Gambar 5.10. Total Distribusi Topik Tiap Kategori Berdasarkan Bulan

Halaman sengaja dikosongkan

BAB VI

HASIL DAN PEMBAHASAN

Bagian ini akan menjelaskan mengenai hasil dan pembahasanyang diperoleh dari penelitian.

6.1.Load Data

Pemuatan data yang dilakukan pada penelitian ini, memiliki jumlah keseluruhan 4400 data. Untuk gambaran dari jumlah data dapat dilihat pada 6.1.

Tabel 6. 1. Jumlah data suara pelanggan

Bulan	2014	2015	2016
Januari	238	189	160
Februari	193	128	148
Maret	159	150	180
April	138	132	199
Mei	112	120	102
Juni	151	130	151
Juli	102	114	-
Agustus	119	128	-
September	145	167	-
Oktober	158	141	-
November	144	132	-
Desember	165	105	-
Jumlah	1824	1636	940
Total	4400		

Data pada bulan Januari 2014 berjumlah 238 data. Data pada bulan Februari 2014 berjumlah 193 data. Data pada bulan Maret 2014 berjumlah 159 data. Data pada bulan April 2014 berjumlah 138 data. Data pada bulan Mei 2014 berjumlah 112 data. Data pada bulan Juni 2014 berjumlah 151 data. Data pada bulan Juli 2014 berjumlah 102 data. Data pada bulan

Agustus 2014 berjumlah 119 data. Data pada bulan September 2014 berjumlah 145 data. Data pada bulan Oktober 2014 berjumlah 158 data. Data pada bulan November 2014 berjumlah 144 data. Data pada bulan Desember 2014 berjumlah 165 data.

Data pada bulan Januari 2015 berjumlah 189 data. Data pada bulan Februari 2015 berjumlah 128 data. Data pada bulan Maret 2015 berjumlah 150 data. Data pada bulan April 2015 berjumlah 132 data. Data pada bulan Mei 2015 berjumlah 120 data. Data pada bulan Juni 2015 berjumlah 130 data. Data pada bulan Juli 2015 berjumlah 114 data. Data pada bulan Agustus 2015 berjumlah 128 data. Data pada bulan September 2015 berjumlah 167 data. Data pada bulan Oktober 2015 berjumlah 141 data. Data pada bulan November 2015 berjumlah 132 data. Data pada bulan Desember 2015 berjumlah 105 data.

Data pada bulan Januari 2016 berjumlah 160 data. Data pada bulan Februari 2016 berjumlah 148 data. Data pada bulan Maret 2016 berjumlah 180 data. Data pada bulan April 2016 berjumlah 199 data. Data pada bulan Mei 2016 berjumlah 102 data. Data pada bulan Juni 2016 berjumlah 151 data.

6.2.Percobaan pra-proses data

Percobaan pra-proses data dilakukan untuk mencari dan menentukan data siap olah dengan kualitas terbaik. Data dalam penelitian digunakan sebagai indikator untuk mengetahui seberapa baik sebuah data jika akan diproses ke dalam tahapan selanjutnya. Percobaan pra-proses dilakukan dengan mengubah nilai-nilai variabel data kemudian membandingkannya dengan hasil data yang tidak disertai dengan pra-proses. Tahapan pra-proses meliputi beberapa langkah diantaranya adalah stemming.

Dalam penelitian ini, dilakukan beberapa skenario percobaan pra proses data untuk mengetahui seberapa penting stemming untuk dilakukan terhadap data yang ada.

- 1.Data dengan proses stemming
- 2.Data tanpa stemming

3.Data hasil stemming dengan pengecekan

Data dengan proses stemming

Hasil dari data dengan stemming merupakan kata-kata yang telah dihilangkan kata imbuhan seperti pe, me, an sehingga menjadi kata baku. Dari hasil stemming yang dilakukan kemudian disimpan ke dalam sebuah file untuk selanjutnya digunakan sebagai korpus pada pemrosesan data.

Data tanpa stemming

Data tanpa stemming merupakan gabungan data yang secara langsung diolah sebagai input pada proses data. Data tidak diubah menjadi kalimat dengan kata baku, namun tetap dengan struktur asli kalimat sesuai yang ada dalam dokumen.

Data hasil stemming dengan pengecekan

Hasil dari data stemming dengan pengecekan merupakan kata-kata hasil dari skenario 1 yang kemudian dilakukan pengecekan kembali terhadap masing-masing kata dalam dokumen. Jika terdapat suatu kata yang memiliki keanehan, maka diubah ke dalam kata dasar yang seharusnya. Sebagai salah satu contohnya adalah kata “pupuk” yang menjadi “pupu” dalam dokumen. Sehingga, kata tersebut harus diubah karena dapat memiliki arti yang berbeda. Kata-kata yang diubah secara manual diantara lain adalah pupu, recruitmen, distrinutor, penanggungjawab, mengujicobakan, medapatkan, mensosialisasikan. Hal yang dapat mempengaruhi hasil stemming ini adalah penulisan dari pengguna aplikasi yang menginputkan form suara pelanggan. Apabila terdapat typo atau kesalahan penulisan, maka kata-kata dalam dokumen akan memiliki perbedaan arti.

Setelah melalui tahapan stemming, data selanjutnya diolah dengan melakukan pendefinisian stopword, yang artinya merupakan penghilangan kata-kata yang sering muncul namun tidak memiliki arti pada dokumen. Kata-kata stopword merupakan kata-kata yang diambil dari kata tugas yang

merujuk pada buku Tata Baku Bahasa Indonesia milik Hasan Alwi. Ketiga hasil dokumen dari skenario sebelumnya, melewati tahapan stopword removal dengan kata-kata yang dapat dilihat pada lampiran A.

Selanjutnya, pada tahapan tokenization, yaitu tahapan yang menghilangkan karakter-karakter tertentu berupa angka dan karakter lainnya yang tidak diperlukan ketika melakukan analisis terhadap dokumen. Data kemudian diproses dengan cara melihat jumlah kemunculan kata pada setiap dokumen. Jika kata memiliki jumlah kemunculan lebih dari 2 kali, maka kata tersebut dianggap unik dan disimpan ke dalam suatu list yang akan digunakan sebagai input pada tahapan pembuatan model.

Ketiga hasil dokumen dari skenario yang digunakan, dievaluasi berdasarkan nilai perplexity yang dihasilkan oleh suatu model secara acak. Hasil perplexity menunjukkan bahwa data dengan stemming disertai dengan pengecekan lebih baik daripada data yang tidak di stem dan data yang di stemming tanpa melalui proses pengecekan kembali. Hasil perbandingan perplexity ketiga data dapat dilihat pada lampiran A.

6.3.Pembuatan Model LDA

Pembuatan model LDA merupakan suatu tahapan dimana data dianalisis untuk mendapatkan model yang menghasilkan output terbaik. Model LDA merupakan inti dari penelitian ini. Model LDA dicari berdasarkan beberapa aspek, diantaranya iterasi dan jumlah topik.

6.3.1. Penentuan jumlah iterasi

Jumlah iterasi merupakan suatu tahapan yang digunakan dalam menentukan seberapa stabil model ketika dijalankan. Stabilitas model dapat diartikan sebagai tingkat kestabilan atau tidak berubahnya nilai perplexity yang dihasilkan. Penentuan jumlah iterasi dilakukan dengan melakukan percobaan sebanyak 3 kali. Hasil perplexity didapatkan dari model awal dengan iterasi sejumlah 30 kali. Selain dari iterasi, percobaan juga dilakukan dengan membandingkan jumlah topik yang

digunakan, yaitu 5 topik, 10 topik dan 15 topik. Hasil berupa tabel uji perplexity model dapat dilihat pada lampiran A.

Berikut merupakan skenario yang digunakan dalam melakukan percobaan penentuan iterasi:

Percobaan 1

1. Percobaan 30 kali iterasi, dengan jumlah topik=5
2. Percobaan 30 kali iterasi, dengan jumlah topik=10
3. Percobaan 30 kali iterasi, dengan jumlah topik=15

Percobaan 2

1. Percobaan 30 kali iterasi, dengan jumlah topik=5
2. Percobaan 30 kali iterasi, dengan jumlah topik=10
3. Percobaan 30 kali iterasi, dengan jumlah topik=15

Percobaan 3

1. Percobaan 30 kali iterasi, dengan jumlah topik=5
2. Percobaan 30 kali iterasi, dengan jumlah topik=10
3. Percobaan 30 kali iterasi, dengan jumlah topik=15

Hasil dari evaluasi perplexity pada setiap percobaan, ditampilkan pada Gambar 6.1., Gambar 6.2., dan Gambar 6.3. Dari hasil Gambar 6.1. dapat diketahui bahwa nilai perplexity yang dihasilkan, semakin menurun pada setiap penambahan iterasi. Hasil perplexity yang mengalami penurunan secara signifikan ditunjukkan pada titik 2. Hal ini menunjukkan bahwa dengan melakukan perulangan pada model, maka model akan semakin belajar atau mengalami proses *learning* kemudian semakin bertambah pintar. Model mulai stabil pada iterasi ke 15, namun dari Gambar 6.1. pada percobaan 1, model masih mengalami peningkatan selisih hasil perplexity pada iterasi ke 17 yaitu sejumlah 0,2.

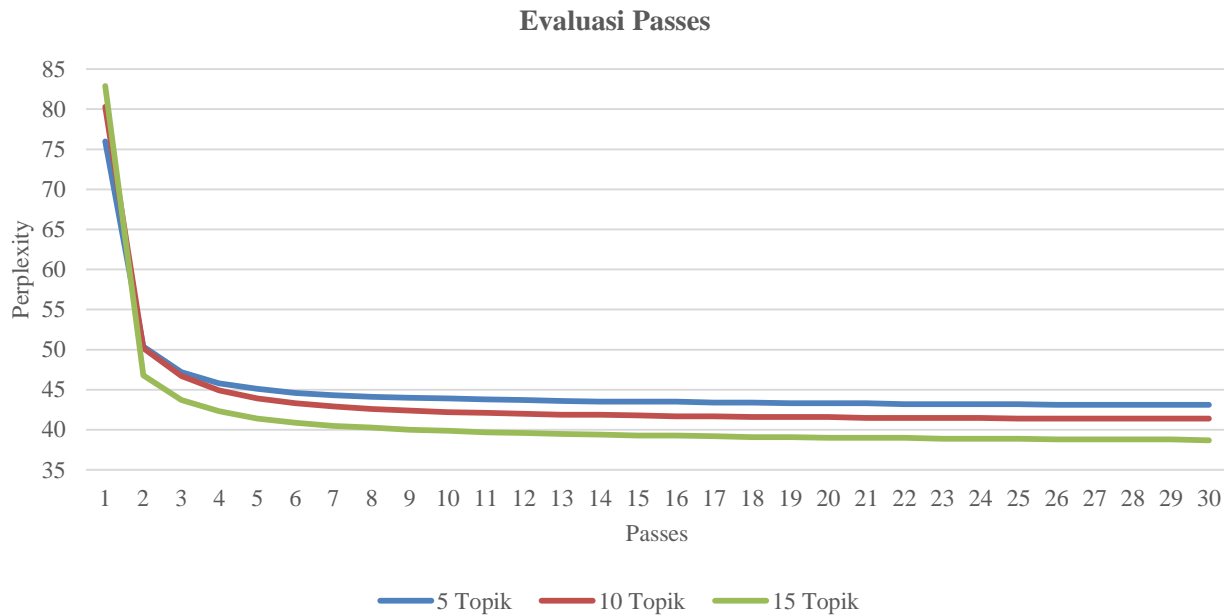
Jika dianalisa lebih lanjut, uji perplexity pada model akan mulai kembali stabil pada iterasi ke 20. Angka tersebut merupakan titik kestabilan model sehingga perplexity yang

dihasilkan cenderung tetap dengan nilai selisih antara 0 sampai 0.1.

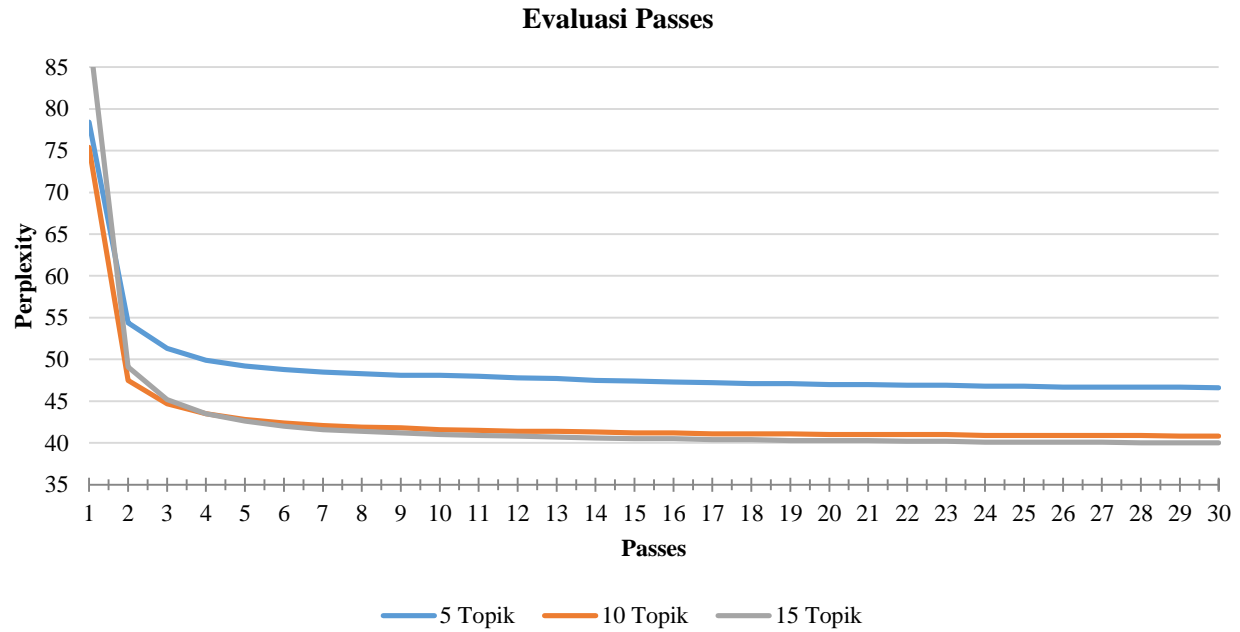
Evaluasi perplexity juga dapat dibandingkan berdasarkan jumlah topik yang digunakan sebagai variabel pengujian. Dengan jumlah topik sama dengan 5, evaluasi perplexity cenderung masih memiliki nilai yang tinggi yaitu sekitar 43,3. Sedangkan dengan jumlah topik sama dengan 10, evaluasi perplexity yang dihasilkan sedikit menurun yakni sekitar 41,6. Evaluasi perplexity dari jumlah topik sama dengan 15, memiliki nilai yang paling rendah yaitu berkisar pada rentang 39.

Dari hasil Gambar 6.2. dapat diketahui bahwa nilai perplexity yang dihasilkan, semakin menurun pada setiap penambahan iterasi. Penurunan hasil perplexity secara signifikan ditunjukkan pada titik 2. Hal ini menunjukkan bahwa dengan melakukan perulangan pada model, maka model akan semakin belajar atau mengalami proses *learning* serta bertambah pintar. Model sudah mengalami titik stabil pada iterasi ke 15 dengan selisih hasil perplexity pada iterasi berkisar antara 0-0.1.

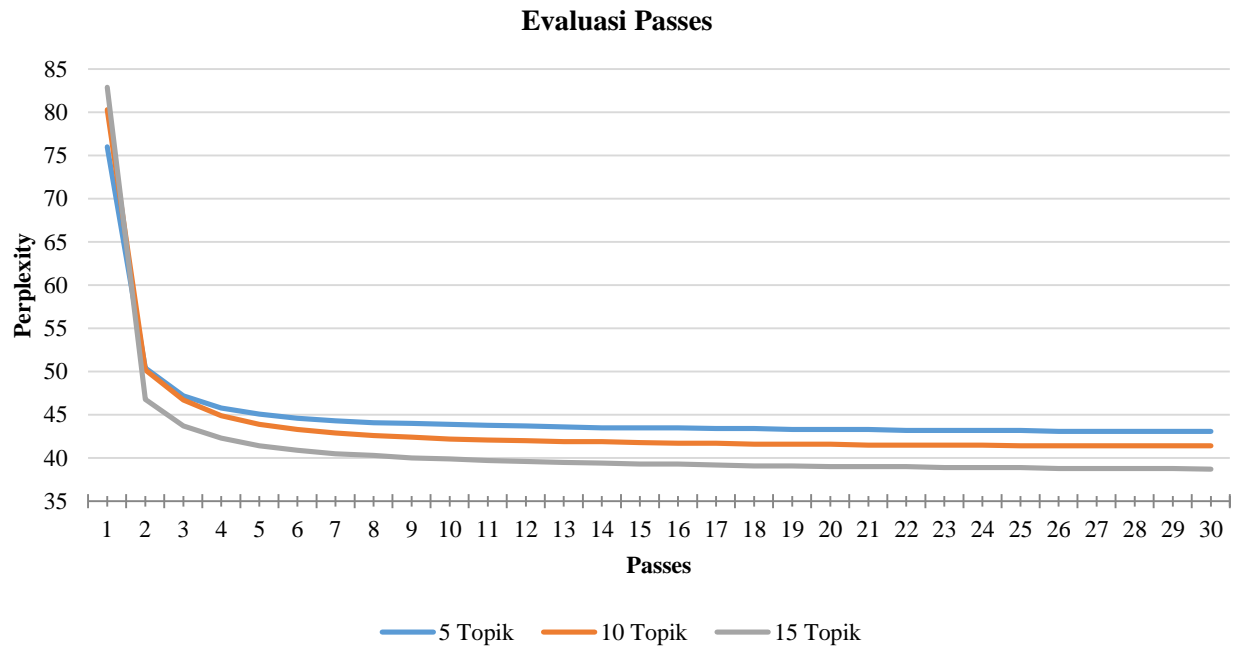
Evaluasi perplexity juga dapat dibandingkan berdasarkan jumlah topik yang digunakan sebagai variabel pengujian. Dengan jumlah topik sama dengan 5, evaluasi perplexity cenderung masih memiliki nilai yang tinggi yaitu sekitar 47. Sedangkan dengan jumlah topik sama dengan 10, evaluasi perplexity yang dihasilkan sedikit menurun yakni sekitar 41. Evaluasi perplexity dari jumlah topik sama dengan 15, memiliki nilai yang paling rendah yaitu berkisar pada rentang 40,3.



Gambar 6. 1. Percobaan 1 Evaluasi Perplexity dengan 30 iterasi



Gambar 6. 2. Percobaan 2 Evaluasi Perplexity dengan 30 iterasi



Gambar 6. 3. Percobaan 3 Evaluasi Perplexity dengan 30 iterasi

Dari hasil Gambar 15 dapat diketahui bahwa nilai perplexity yang dihasilkan, semakin menurun pada setiap penambahan iterasi. Hasil perplexity yang mengalami penurunan secara signifikan ditunjukkan pada titik 2. Model mulai stabil pada iterasi ke 15, namun dari Gambar 15 pada percobaan 3, model masih mengalami peningkatan selisih hasil perplexity pada iterasi ke 16 yaitu dari 0,1 menjadi 0,2.

Setelah dianalisa secara lebih lanjut, model akan mulai kembali stabil pada iterasi ke 20. Angka tersebut merupakan titik kestabilan model sehingga perplexity yang dihasilkan cenderung tetap dengan nilai selisih antara 0 sampai 0.1.

Evaluasi perplexity juga dapat dibandingkan berdasarkan jumlah topik yang digunakan sebagai variabel pengujian. Dengan jumlah topik sama dengan 5, evaluasi perplexity cenderung masih memiliki nilai yang tinggi yaitu sekitar 43,3. Sedangkan dengan jumlah topik sama dengan 10, evaluasi perplexity yang dihasilkan yakni sekitar 41,6. Evaluasi perplexity dari jumlah topik sama dengan 15, memiliki nilai yang paling rendah yaitu berkisar pada rentang 39.

6.3.2. Penentuan jumlah topik

Penentuan jumlah topik dapat diukur untuk menggambarkan seberapa tepat dokumen baru jika akan dikelompokkan kedalam topik tersebut. Semakin sedikit jumlah topik menunjukkan bahwa topik terlalu general dan saling bercampur antara bahasan satu dengan yang lainnya. Semakin banyak jumlah topik menunjukkan terlalu spesifik topik sehingga mengandung kata-kata yang tidak dapat disimpulkan. Rentang pemilihan topik dalam penelitian ini adalah antara 3-50 topik. Dengan angka pada rentang tersebut, penelitian ini juga menganalisa apakah dengan jumlah topik tertentu, setiap topik yang muncul memiliki hirarki tersendiri. Selain itu, topik juga dapat dianalisis apakah kata-kata yang terdapat dalam topik justru memiliki kesamaan antara 1 topik dengan topik yang lain, sehingga dapat digabungkan. Setelah dilakukan percobaan terhadap data, jumlah topik yang memiliki elbow paling tinggi adalah jumlah topik 30, 35, 40 dan 45. Oleh

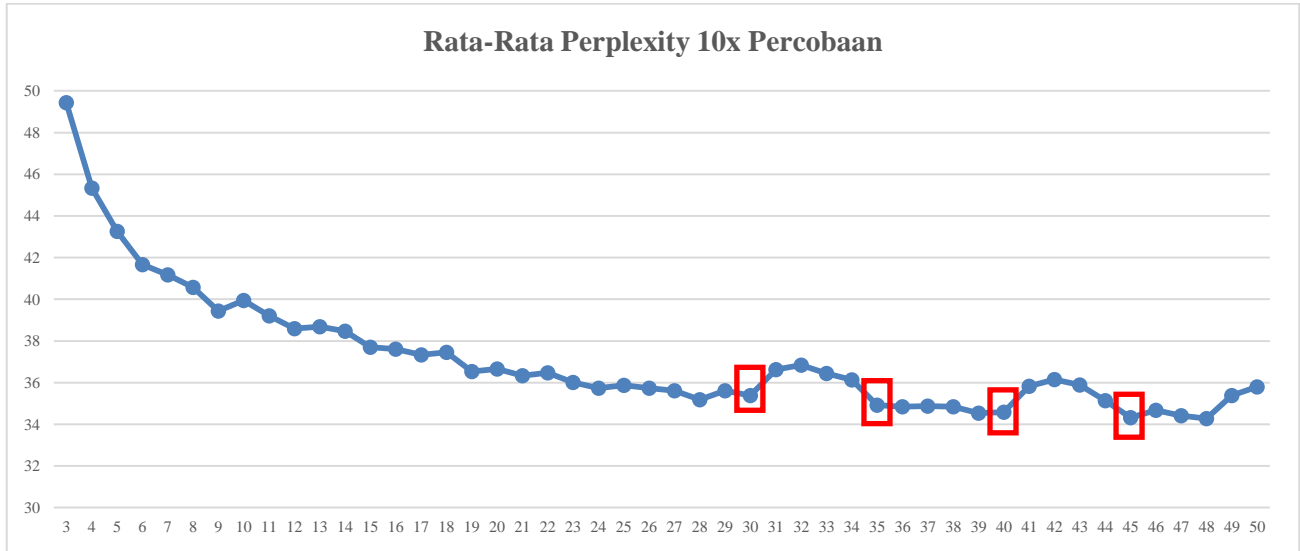
karena itu selanjutnya dilakukan analisis lebih lanjut terhadap keempat jumlah topik tersebut. Hasil rata-rata perplexity pada setiap topik ditampilkan pada gambar 16.

Kata-kata yang terdapat dalam setiap topik yaitu 30, 35, 40 dan 45 ditunjukkan dalam tabel. Kata-kata dalam topik menunjukkan bahwa dalam setiap topik, mengandung probabilitas terhadap kata tersebut sebagai kata yang sering muncul dalam dokumen.

Kata dievaluasi berdasarkan koherensi dengan kata yang lain yang muncul dalam satu topik. Jika kata-kata yang dihasilkan semakin terpusat, maka topik yang dihasilkan semakin bagus. Sedangkan jika kata-kata dalam satu topik tidak terlalu koheren, maka topik tidak dapat disimpulkan berdasarkan kata-kata yang tergantung di dalam topik.

Evaluasi kata-kata dalam setiap topik yang muncul digunakan untuk mengetahui jumlah topik yang akan digunakan sebagai input untuk melakukan analisis topik serta penamaan topik pada tahapan selanjutnya.

Tabel xx menunjukkan distribusi kata-kata yang terdapat dalam topik dengan jumlah topik 30. Dari hasil yang ditunjukkan, terdapat beberapa topik yang mengandung campuran kata yang terlalu general seperti pada topik 8 yaitu ana, hari, bondowoso, jual, memorandum, juni, het, atas, tebus, za. Topik 8 menunjukkan distribusi topik antara pemberitaan, harga, dan penjualan. Sedangkan pada jumlah topik 12 dengan distribusi kata-kata yaitu konsumen, kepada, laksana, rekomendasi, sosialisasi, tani, imbang, jelas, sahabat, kirim juga mengandung beberapa kata-kata yang terlalu general sehingga tidak dapat disimpulkan untuk menjadi topik tertentu.



Gambar 6. 4. Rata-Rata Perplexity 50 Topik

Kata-kata yang terkandung dalam topik 30 merupakan kata-kata hasil luaran proses dari model LDA yang ditampilkan pada Tabel 6.2.

Tabel 6. 2. Kata-kata pada topik 30

Topik	Kata-kata dalam topik
1	gresik, petrokimia, pt, perlu, masalah, positif, tahun, cegah, tidak, produksi
2	pupuk, tanam, tanah, petro, konsumen, cara, aplikasi, biofertil, jagung, tingkat
3	konsumen, petrokimia, gresik, pt, arah, produk, departemen, resmi, syarat, beli
4	gresik, petrokimia, pt, konsumen, informasi, recruitment, karyawan, dapat, tidak, laku
	pupuk, kg, dosis, konsumen, phonska, urea, padi, pohon, hst, petroganik
6	pupuk, npk, kebomas, nonsubsidi, konsumen, harga, beli, ton, pesan, korporasi
7	konsumen, layan, pt, petrosida, coba, beli, produk, petro, wilayah, pasar
8	ana, hari, bondowoso, jual, memorandum, juni, het, atas, tebus, za
9	hi, corn, nasional, wajib, merk, indonesia, koordinator, butuh, bunga, brebes
10	life, asuransi, pensiun, lampung, insurance, pkbl, kerja, kaptan, program, billing
11	pupuk, subsidi, tani, ton, tidak, stok, sebut, berita, ada, aman
12	konsumen, kepada, laksana, rekomendasi, sosialisasi, tani, imbang, jelas, sahabat, kirim
13	pupuk, subsidi, rdkk, konsumen, beli, kelompok, butuh, cara, masuk, kios

Topik	Kata-kata dalam topik
14	milik, distributor, konsumen, syarat, kantor, transportasi, badan, kurang, pupuk, subsidi
15	gresik, petrokimia, pt, program, tani, bina, dukung, positif, masalah, cegah
16	gresik, bantu, pg, petrokimia, usaha, giat, sekitar, pt, kepada, festival
17	pupuk, phonska, konsumen, kantong, npk, kemas, produksi, tidak, produk, unsur
18	pupuk, pt, konsumen, petrokanik, wilayah, indonesia, ada, hubung, investor, jelas
19	bulan, harga, petrokimia, gresik, murah, usaha, masjid, bumn, lama, ramadhan
20	pasar, komoditas, bahan, paket, masing, sekretaris, wahyudi, baku, rabu, lokal
21	pupuk, kios, tani, laku, kecamatan, ton, ss, sp, distributor, sebut
22	benih, padi, petroseed, rock, phosphate, tani, varietas, kapur, lahan, konsumen
23	dapat, nomor, tambak, telepon, distributor, kpg, sulit, gresik, beli, rdck
24	usaha, modal, kerja, pinjam, miliar, pg, mitra, salur, kembang, jawa
25	kios, resmi, konsumen, distributor, pupuk, subsidi, sebut, syarat, aju, tanggung
26	gas, standar, suprapto, telah, ponorogo, pertama, milik, nganjuk, bojonegoro, sebut
27	petrokimia, gresik, inovasi, tingkat, giat, usaha, pt, nugroho, christijanto, becak
28	pt, tani, obat, petrokimia, kaya, tanam, ext, sakit, pardiman, gresik

Topik	Kata-kata dalam topik
29	riset, hipa, hybrid, hasil, kategori, awards, anugerah, lantasi, agenda, laut
30	pabrik, dana, petrokimia, pt, gresik, bni, amoniak, pkg, bangun, global

Kata-kata yang terdapat dalam topik 35 merupakan kata-kata hasil luaran proses dari model LDA yang ditampilkan pada Tabel 6.3.

Tabel 6. 3. Kata-kata pada topik 35

Topik	Kata-kata dalam topik
1	pupuk, phonska, indonesia, konsumen, tidak, produksi, npk, petroganik, kantong, tahu
2	konsumen, gresik, pupuk, produk, petrokimia, pt, coba, nonsubsidi, arah, resmi
3	bni, life, asuransi, kelola, karyawan, pensiun, dana, kerjasama, insurance, kerja
4	inovasi, harga, hemat, pasar, tingkat, murah, terus, budaya, kpg, giat
5	pupuk, harga, konsumen, npk, beli, kebomas, nonsubsidi, pesan, ton, korporasi
6	gresik, petrokimia, pt, pg, usaha, giat, tidak, nugroho, berita, masalah
7	gresik, petrokimia, pt, tani, tingkat, ukm, pangan, program, perlu, masalah
8	gresik, petrokimia, pt, usaha, perlu, positif, cegah, masalah, berita, pabrik
9	via, tani, kaptan, pupuk, pt, ton, wonogiri, coba, petrokimia, kapur
10	bantu, benih, ramadhan, bulan, warga, sekitar, kepada, petroseed, dapat, bina
11	konsumen, departemen, petro, hipa, hybrid, potensi, pasar, wilayah, informasi, beli

Topik	Kata-kata dalam topik
12	milik, distributor, konsumen, syarat, resmi, badan, kantor, gresik, sarana, jelas
13	pupuk, konsumen, dosis, padi, phonska, kg, urea, hst, area, base
14	tani, tabloid, sahabat, konsumen, kirim, alamat, terima, kasih, atas, rasa
15	petro, konsumen, tambak, biofertil, promosi, kirim, fish, lombok, koordinasi, ikan
16	laku, produk, wujud, dasar, saat, hadap, hasil, proses, muat, kantong
17	pohon, pupuk, kg, konsumen, dosis, tanam, sawit, kelapa, sumatera, belum
18	festival, asuh, panti, salur, pesantren, pertagas, pondok, milik, pertamina, lantasi
19	manyar, kunjung, pupuk, duga, bahan, baku, kp, malang, sulfat, kasus
20	kios, resmi, konsumen, distributor, subsidi, syarat, pupuk, aju, sebut, punya
21	kepada, laksana, rekomendasi, sosialisasi, imbang, semester, phonska, pokok, urea, petrogranik
22	pupuk, gresik, ton, subsidi, petrokimia, pt, tani, stok, salur, alokasi
23	pupuk, subsidi, sak, gas, het, kios, harga, resmi, konsumen, beli
24	kg, za, komoditas, pupuk, cabai, jagung, cara, aplikasi, produktif, tanam
25	humas, departemen, laku, koordinasi, jatim, belum, tahu, tender, pupuk, pasu
26	konsumen, layan, pt, petrosida, gresik, beli, pasar, distributor, bebas, sms
27	pupuk, timur, wilayah, jawa, sistem, pt, kalimantan, surabaya, barat, organik
28	konsumen, pt, npk, kebomas, pupuk, hubung, arah, ext, departemen, beli

Topik	Kata-kata dalam topik
29	riset, fitrice, produk, guna, konsumen, rice, fit, rpph, sponsorship, swalayan
30	petrokimia, gresik, pt, konsumen, recruitment, informasi, dapat, karyawan, kerja, laku
31	ana, atas, tata, cara, aku, lancar, hubung, ternak, jabar, pt
32	pupuk, kios, tani, kecamatan, distributor, sebut, dapat, kabupaten, desa, tidak
33	ada, bojonegoro, lamongan, modal, lampung, distribusi, tuban, billing, mudah, nganjuk
34	pupuk, subsidi, rdkk, konsumen, beli, kelompok, butuh, cara, masuk, dapat
35	petroganik, tanah, pupuk, konsumen, manfaat, tanam, jelas, tarik, produk, efektivitas

Kata-kata yang terdapat dalam topik 40 merupakan kata-kata hasil luaran proses dari model LDA yang ditampilkan pada Tabel 6.4.

Tabel 6. 4. Kata-kata pada topik 40

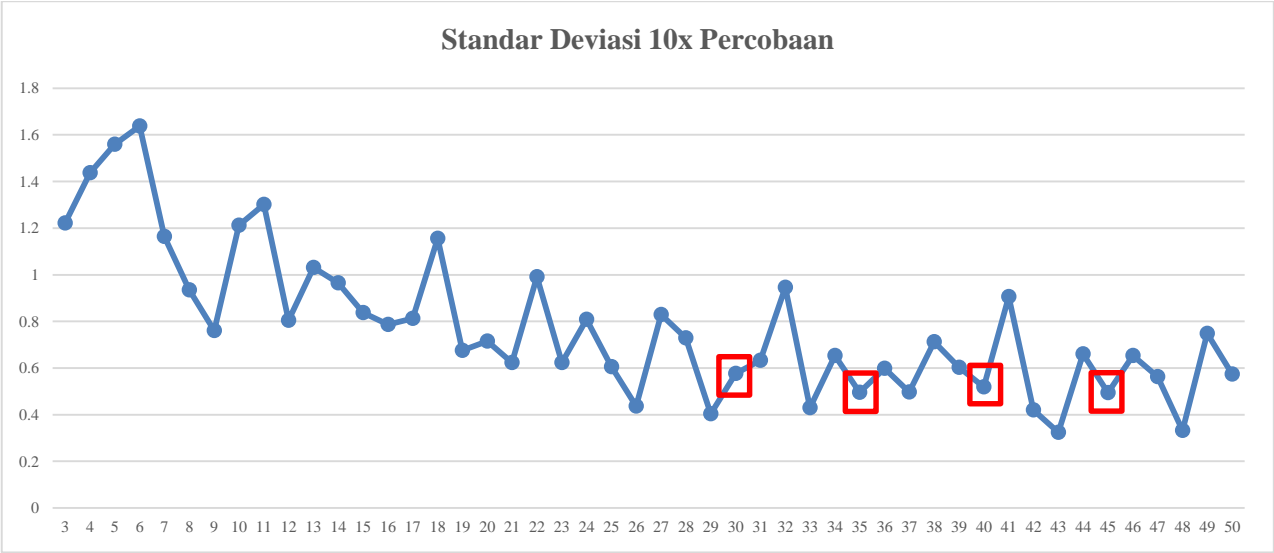
Topik	Kata-kata dalam topik
1	pohon, kg, pupuk, laksana, dosis, konsumen, sawit, kelapa, phonska, tanam
2	nomor, telepon, cv, produk, langsung, swalayan, gladiator, petik, kulonprogo, cilacap
3	gresik, petrokimia, usaha, pt, bantu, positif, tidak, berita, cegah, perlu
4	petrokimia, kaya, didik, ext, keluarga, pt, bio, dapat, kodim, pasar
6	gresik, petrokimia, pt, dana, pabrik, bangun, bni, perlu, cegah, positif
7	pupuk, kg, konsumen, dosis, urea, padi, phonska,

Topik	Kata-kata dalam topik
	hst, imbang, area
8	harga, murah, pokok, awards, gelar, stevie, pasific, asia, pertama, bahan
9	pg, saing, hemat, modal, baik, laku, pinjam, global, usaha, kuliner
10	petro, tanah, tanam, konsumen, produk, pupuk, biofertil, baik, jelas, petroganik
11	pupuk, phonska, npk, kantong, konsumen, kemas, unsur, tidak, indonesia, cantum
12	konsumen, layan, petro, petrosida, pt, produk, sms, distributor, bebas, pulsa
13	humas, koordinasi, laku, departemen, konsumen, prosedur, beli, cara, belum, hubung
14	pupuk, konsumen, nonsubsidi, harga, npk, kebomas, beli, pesan, ton, korporasi
15	pupuk, subsidi, konsumen, rdkk, beli, kelompok, butuh, cara, masuk, jelas
16	milik, distributor, konsumen, syarat, kantor, resmi, transportasi, urus, punya, bank
17	kios, resmi, konsumen, aju, distributor, subsidi, pupuk, syarat, sebut, jawab
18	mail, coba, via, pupuk, pt, konsumen, petroganik, wilayah, indonesia, fitrice
19	pupuk, hari, ana, wujud, milik, jual, bondowoso, uang, atas, kabupaten
20	muat, demplot, bongkar, jiwa, pilih, ton, tang, tua, selain, Palembang
21	pupuk, petrokimia, gresik, pt, tidak, masalah, tani, cegah, perlu, positif
22	festival, culinary, pertagas, hari, Pertamina,

Topik	Kata-kata dalam topik
	tahun, milik, seni, mmscfd, agenda
23	rekomendasi, pt, gresik, tani, merk, beri, kepada, korban, lahan, petrokimia
24	pupuk, subsidi, sak, kios, harga, resmi, konsumen, het, sp, phonska
25	life, asuransi, pensiun, karyawan, tani, program, insurance, negeri, kerja, lombok
26	gas, paket, pkg, ubah, per, standar, pabrik, salur, banyak, telah
27	gresik, petrokimia, pt, konsumen, recruitment, informasi, dapat, karyawan, kerja, laku
28	konsumen, pt, layan, beli, petrosida, gresik, pasar, informasi, potensi, wilayah
29	tani, konsumen, tabloid, sahabat, kirim, suprapto, alamat, rasa, pos, kasih
30	petrokimia, gresik, pg, inovasi, giat, pt, tingkat, terus, positif, usaha
31	ngawi, truk, atas, kami, lantas, sudah, alamat, konsumsi, toko, terima
32	pupuk, laku, sebut, gudang, saat, ss, hasil, tidak, ada, awas
33	konsumen, departemen, wilayah, petro, informasi, hipa, benih, hybrid, tarik, pasar
34	pupuk, tani, kios, distributor, dapat, kecamatan, tidak, subsidi, sebut, desa
35	hkti, kota, tegal, kering, pernah, buah, putar, open, mutu, hijau
36	pupuk, pt, bahan, baku, konsumen, organik, petroganik, produksi, hubung, fosfat
37	tugas, surat, promosi, kirim, tindak, sebut, barang, acara, lanjut, alamat

Topik	Kata-kata dalam topik
38	gresik, program, petrokimia, pt, padi, tani, tingkat, pangan, dukung, hasil
39	ton, pupuk, alokasi, subsidi, salur, sudah, juta, langka, urea, tani
	petrokimia, pt, gresik, produksi, ada, tidak, pihak, berita, sebut, pabrik
40	konsumen, petrokimia, gresik, pt, arah, produk, coba, resmi, riset, syarat

Selain menggunakan koherensi topik, hal lain yang juga dapat dianalisis adalah mengenai standar deviasi yang ada dalam perplexity. Standar deviasi dari perplexity dari keempat topik yang dianalisis memiliki nilai yang cukup rendah. Jumlah topik 30 memiliki nilai standar deviasi 0.58. Sedangkan jumlah topik 35 memiliki nilai standar deviasi 0.5. Jumlah topik 40 memiliki nilai standar deviasi 0.52 dan jumlah topik 45 memiliki nilai standar deviasi 0.49. Standar deviasi menunjukkan simpangan perplexity dari setiap topik. Semakin kecil nilai standar deviasi, maka perplexity dari suatu topik cenderung memiliki angka tetap atau tidak terlalu naik turun. Sedangkan semakin tinggi nilai perplexity menunjukkan bahwa model tidak stabil karena memiliki nilai perplexity yang naik turun.



Gambar 6. 5. Standar Deviasi Perplexity 50 Topik

6.4. Analisis Topik

Dari keempat topik yang dianalisis, topik dievaluasi berdasarkan keterkaitan dengan panduan topik yang ada pada PT. Petrokimia Gresik. Panduan topik dari PT. Petrokimia Gresik adalah sebagai berikut:

1. Produk

- a. Keluhan warna
- b. Keluhan bentuk fisik (keras, lembek, dsb)
- c. Keluhan adanya bahan yang ikut tercampur
- d. Keluhan bantuan
- e. Keluhan tidak sesuai standar spesifikasi
- f. Keluhan akan tingkat kelarutan
- g. Keluhan karung rusak (kena gancu, rapuh, dsb)

2. Harga

Harga pupuk di tingkat kios resmi diatas HET

3. Pembelian

- a. Pemesanan pupuk nonsubsidi.
- b. Pemesanan produk nonpupuk.

4. Stock/Distribusi

- a. Keluhan pupuk di Distributor dan Kios resmi kosong
- b. Kekosongan stock pupuk karena peralihan stok antar gudang penyangga

5. Promosi

- a. Permohonan demplot
- b. Permintaan tabloid sahabat petani dan brosur
- c. Permintaan kaos promosi
- d. Undian berhadiah

6. Pelayanan

- a. Keluhan kurang timbang (OKT)
- b. Keluhan keterlambatan kedatangan pupuk (kelangkaan)
- c. Penjelasan dosis pemupukan.
- d. Permintaan studi lapang
- e. Penjelasan nonproduk PT Petrokimia Gresik.
- f. Menanyakan syarat menjadi distributor dan kios resmi.
- g. Menanyakan lowongan pekerjaan.

7. Lainnya

- a. Ucapan terima kasih dari instansi/perseorangan.
- b. Pengajuan proposal ke PT. Petrokimia Gresik
- c. Permohonan kerjasama
- d. Permohonan kunjungan

Berdasarkan uraian kata-kata yang ada pada topik 30, 35, 40, dan 45, topik yang memiliki kesesuaian adalah topik 35.

Topik 30 memiliki beberapa kata yang terlalu general sehingga dapat membuat topik menjadi tercampur antara satu dengan yang lainnya. Sedangkan topik 40 dan 45 memiliki kata-kata yang terlalu spesifik sehingga membuat topik menjadi tidak bisa disimpulkan

Dari hasil jumlah topik sama dengan 35, maka jika dipetakan ke dalam hasil analisis topik berdasarkan panduan adalah sebagai berikut :

Produk

[Manfaat Produk](#) [1], Kemasan Produk [13], Warna Pupuk [21], Kantong Pupuk [25], [Penggunaan produk](#) [27]

Harga

Harga pupuk [8]

Pembelian

Info pembelian produk : [Pupuk Subsidi](#) [4], Pupuk Non Subsidi [7], Non Pupuk [3, 12]

Stock/Distribusi

[Produksi](#) [11], Permasalahan Distribusi Produk [17], Stock Produk [18, 31]

Promosi

Tabloid Sahabat Petani [5]

Pelayanan

Penjelasan non produk : [penjelasan seputar perusahaan \[2\]](#),
sosialisasi berimbang [22], alamat distributor [26]

[Lowongan pekerjaan \[20\]](#)

Syarat menjadi distributor [28]

Syarat menjadi kios resmi [33]

Dosis Pemupukan [9, 16, 34]

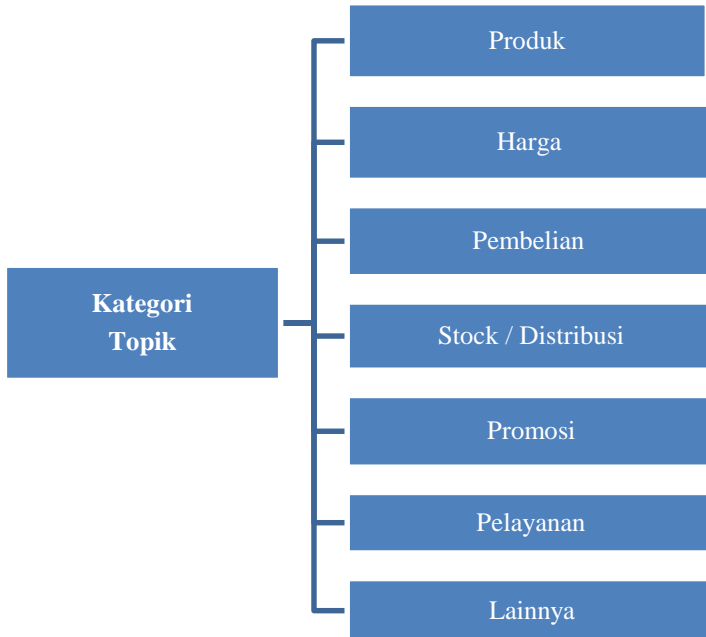
Lainnya

[Kerjasama : kualitas perusahaan \[23, 29\], perjanjian \[30\]](#)

[Permohonan kunjungan dan demplot \[32\]](#)

[Pemberitaan : Kabar positif perusahaan \[0, 14, 15\], kegiatan perusahaan \[6, 19, 24\], penanganan masalah \[10\]](#)

Warna biru dalam kategorisasi menunjukkan kategorisasi baru yang muncul berdasarkan topik yang terbentuk. Jika digambarkan secara hirarkis, maka akan memiliki gambar seperti berikut :



Gambar 6. 6. Kategori Topik

Kategori topik didefinisikan kembali seperti menurut panduan perusahaan yang berjumlah 7 kategori. Berikut merupakan penjelasan isi masing-masing kategori hasil panduan disesuaikan dengan analisis topik yang muncul dalam dokumen suara pelanggan, yaitu :

1. Produk

Kategori produk berisi mengenai topik-topik yang membahas baik berupa pertanyaan, keluhan atau informasi seputar produk secara spesifik.

2. Harga

Kategori harga berisi mengenai topik-topik yang membahas harga, biasanya adalah seputar pertanyaan mengenai harga eceran tertinggi produk.

3. Pembelian

Kategori pembelian berisi mengenai topik-topik yang membahas cara pembelian terhadap produk baik pupuk maupun non pupuk. Biasanya disampaikan melalui pertanyaan.

4. Stock/Distribusi

Kategori Stock/Distribusi membahas mengenai topik-topik tentang kelangkaan produk yang disebabkan oleh kekosongan stock, produksi maupun distribusi yang tersendat. Biasanya disampaikan dengan keluhan, informasi ataupun pertanyaan.

5. Promosi

Kategori promosi membahas mengenai topik-topik yang seputar promosi yaitu tabloid sahabat petani. Biasanya disampaikan melalui informasi.

6. Pelayanan

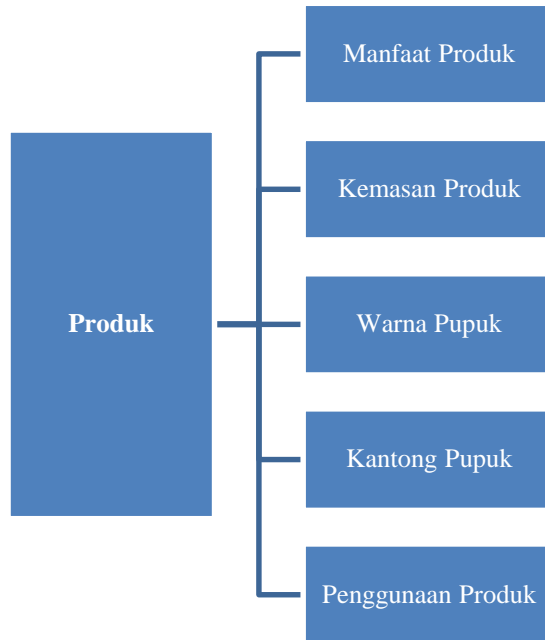
Kategori pelayanan berisi mengenai topik-topik seputar pelayanan kepada pelanggan yaitu berupa pertanyaan seputar dosis pemupukan, lowongan pekerjaan, penjelasan seputar non produk serta penjelasan seputar syarat/cara menjadi distributor dan kios resmi.

7. Lainnya

Kategori lainnya berisi topik-topik seputar pemberitaan serta pengajuan kerjasama terhadap perusahaan. Biasanya disampaikan dalam bentuk laporan berupa informasi hasil pemberitaan ataupun pertanyaan seputar kerjasama dengan perusahaan.

Untuk penjelasan secara detail mengenai isi dari kategori topik, akan dibahas pada bagian berikut:

1. Kategori Produk



Gambar 6. 7. Kategori Topik Produk

Kategori produk meliputi topik-topik seputar manfaat produk, kemasan produk, warna pupuk, kantong pupuk, serta penggunaan produk. Manfaat produk berisi mengenai pertanyaan tentang manfaat dari suatu produk tertentu secara spesifik yang sering ditanyakan oleh pelanggan. Kemasan produk berisi mengenai keluhan pelanggan tentang sesuatu yang terjadi dalam kemasan produk. Seringkali disampaikan dalam bentuk keluhan. Begitu pula dengan warna pupuk dan kantong pupuk. Seringkali warna pupuk dan kantong pupuk merupakan keluhan yang disampaikan oleh pelanggan.

Sedangkan penggunaan produk merupakan penjelasan mengenai cara-cara dalam menggunakan produk. Biasanya disampaikan dalam bentuk pertanyaan.

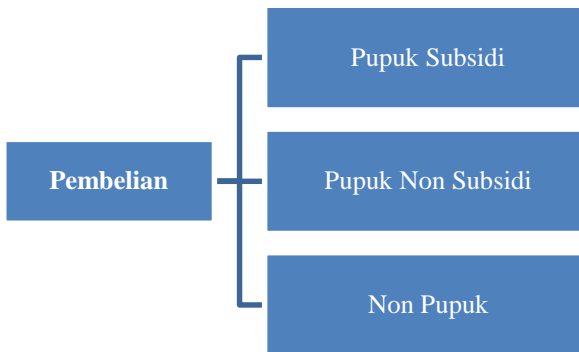
2. Kategori Harga



Gambar 6. 8. Kategori Topik Harga

Kategori harga meliputi pertanyaan atau informasi terkait dengan harga produk yang sering disuarakan oleh pelanggan.

3. Kategori Pembelian

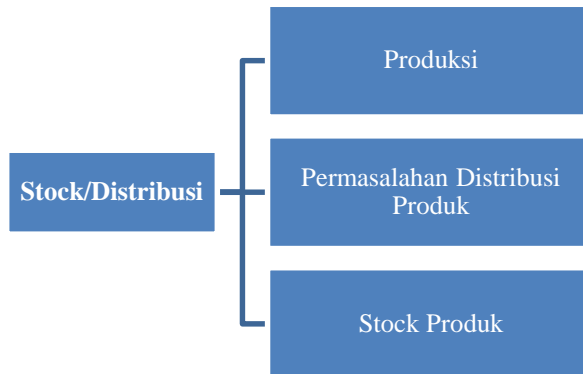


Gambar 6. 9. Kategori Topik Pembelian

Kategori topik pembelian meliputi topik pembelian pupuk subsidi, pembelian pupuk non subsidi, serta pembelian non pupuk. Topik pembelian pupuk subsidi merupakan topik yang

membahas seputar pertanyaan atau informasi pelanggan mengenai cara pembelian pupuk subsidi, rdck, maupun kelompok tani. Sedangkan topik pembelian pupuk non subsidi merupakan topik yang membahas seputar bagaimana cara pembelian pupuk non susidi karena ketertarikan konsumen terhadap produk. Topik Non Pupuk membahas mengenai cara pembelian produk yang bersifat non pupuk seperti fit rice, benih hybrid serta produk pengembangan lain oleh perusahaan.

4. Kategori Stock/Distribusi



Gambar 6. 10. Kategori Topik Stock/Distribusi

Kategori topik stock/distribusi meliputi topik produksi, permasalahan distribusi produk serta stock produk. Topik produksi merupakan topik yang sering diterima dengan jenis informasi dari suara pelanggan. Topik produksi terkategori ke dalam stock/distribusi karena jumlah produksi perusahaan sangat mempengaruhi jumlah stock dan distribusi. Topik permasalahan distribusi produk merupakan topik mengenai keterlambatan distribusi, distribusi yang tidak

mencukupi, serta beberapa permasalahan distribusi lainnya. Sedangkan topik stock produk merupakan hal-hal mengenai kekosongan stock, stock yang terlalu menumpuk, stock yang tidak dapat memenuhi permintaan.

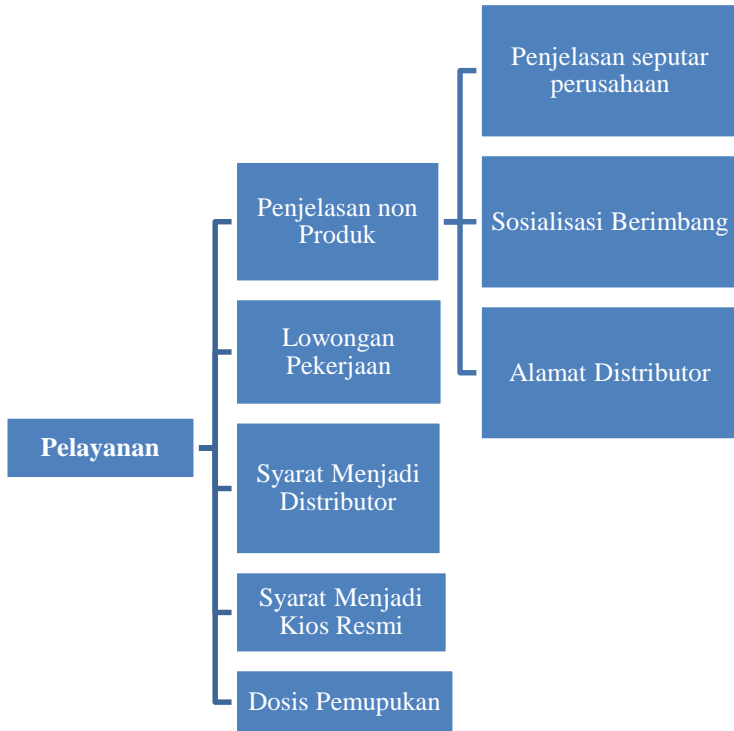
5. Kategori Promosi



Gambar 6. 11. Kategori Topik Promosi

Kategori promosi merupakan kategori yang membahas topik seputar tabloid sahabat tani. Tabloid sahabat tani merupakan salah satu tools atau alat yang digunakan perusahaan dalam melakukan promosi. Oleh karena itu, segala hal yang berhubungan dengan tabloid sahabat tani merupakan topik-topik mengenai promosi.

6. Kategori Pelayanan



Gambar 6. 12. Kategori Topik Pelayanan

Kategori pelayanan merupakan kategori yang membahas topik seputar lowongan pekerjaan, syarat menjadi distributor, syarat menjadi kios resmi, dosis pemupukan serta sub kategori berupa penjelasan non produk. Sub kategori penjelasan non produk merupakan sub kategori yang membahas hal-hal seputar perusahaan, sosialisasi pemupukan berimbang, ataupun penjelasan alamat distributor.

Topik lowongan pekerjaan merupakan topik yang sering ditanyakan oleh pelanggan terkait rekrutmen serta kekosongan posisi pada perusahaan. Topik syarat menjadi distributor dan cara menjadi kios resmi merupakan topik-topik

yang sering dibahas pelanggan dengan tujuan dapat bekerjasama dengan perusahaan. Topik dosis pemupukan merupakan topik yang sering ditanyakan oleh pelanggan terkait dengan takaran dosis serta cara penggunaan maupun aplikasi dari produk.

7. Kategori Lainnya



Gambar 6. 13. Kategori Topik Lainnya

Kategori topik lainnya merupakan kategori yang membahas topik-topik terkait pemberitaan, kerjasama dan permohonan kunjungan dan demplot. Kerjasama merupakan sub kategori dari topik kualitas perusahaan serta perjanjian. Sedangkan pemberitaan merupakan sub kategori dari topik kabar positif perusahaan, kegiatan perusahaan, dan kegiatan masalah. Sub

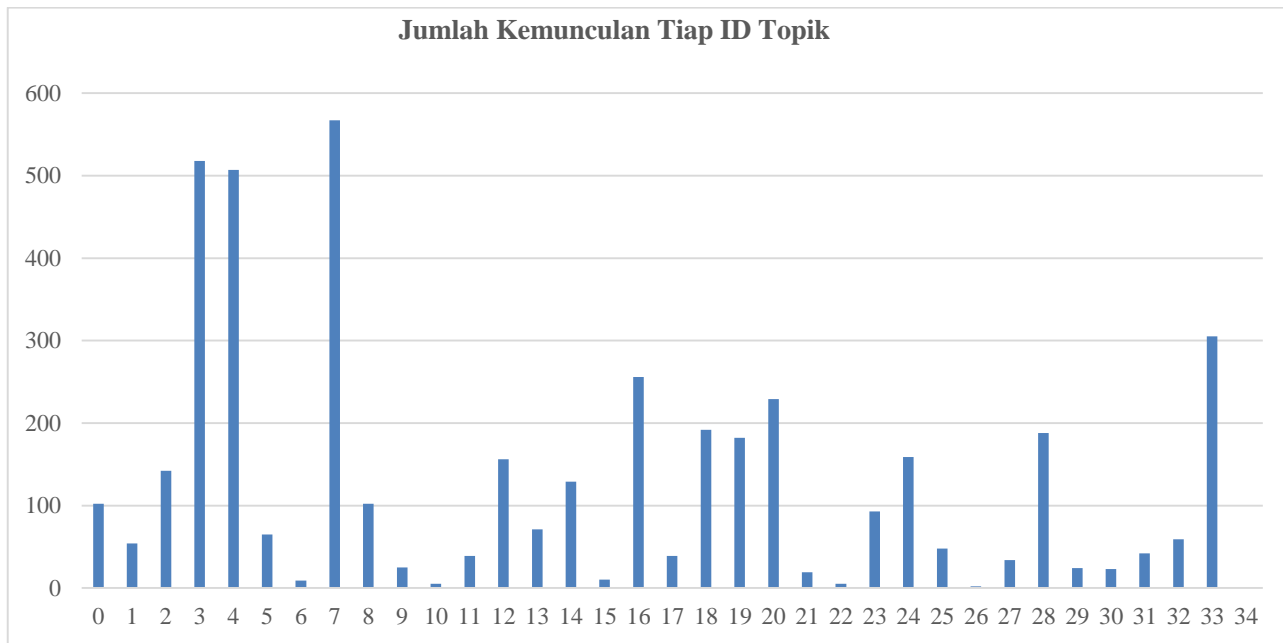
kategori pemberitaan menjelaskan mengenai berita-berita yang dimuat pada media yang menyangkut terhadap nama perusahaan.

6.5.Klasifikasi Data

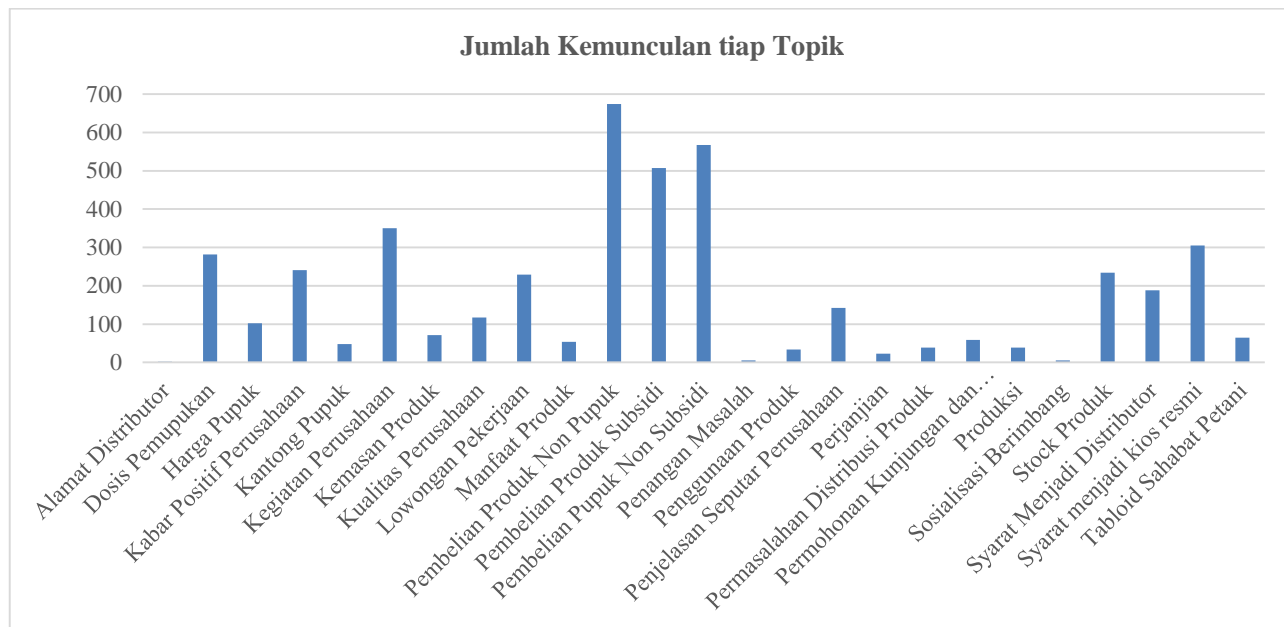
Klasifikasi data merupakan pelabelan data yang dilakukan terhadap dokumen dengan topik yang telah ditentukan. Klasifikasi data dilakukan berdasarkan nilai maksimum setiap probabilitas topik pada suatu dokumen. Klasifikasi data dilakukan pada 4400 dokumen. Kesalahan klasifikasi data bisa terjadi diakibatkan oleh kata-kata yang ada dalam dokumen yang terlalu banyak campuran topik sehingga proporsi terhadap topik yang paling besar belum tentu merepresentasikan dokumen tersebut. Selain itu, terdapat kemungkinan terjadinya beberapa pra-pemrosesan data yang membuat kata-kata dalam dokumen tidak teridentifikasi.

Hasil dari klasifikasi data, Gambar 6.14 menunjukkan bahwa distribusi dokumen pada setiap ID memiliki jumlah paling besar pada ID 7 dengan total dokumen yaitu 567. Selanjutnya merupakan ID 3 sejumlah 517 dokumen dan ID 4 sejumlah 507 dokumen. ID 7 merupakan topik pembelian produk nonsubsidi, sedangkan ID 3 merupakan topik pembelian produk non pupuk serta ID 4 merupakan topik pembelian produk subsidi.

Pada penelitian ini, setelah dilakukan analisis topik, karena terdapat beberapa ID yang memiliki kesamaan arti topik, maka pada Gambar 6.15 menunjukkan distribusi kata pada setiap topik dalam dokumen. 3 topik yang memiliki peran dominan adalah topik pembelian non produk, pembelian produk nonsubsidi, dan pembelian produk subsidi, dengan masing-masing jumlah kemunculan 674, 567, dan 507 dokumen.



Gambar 6. 14. Distribusi topik setiap ID

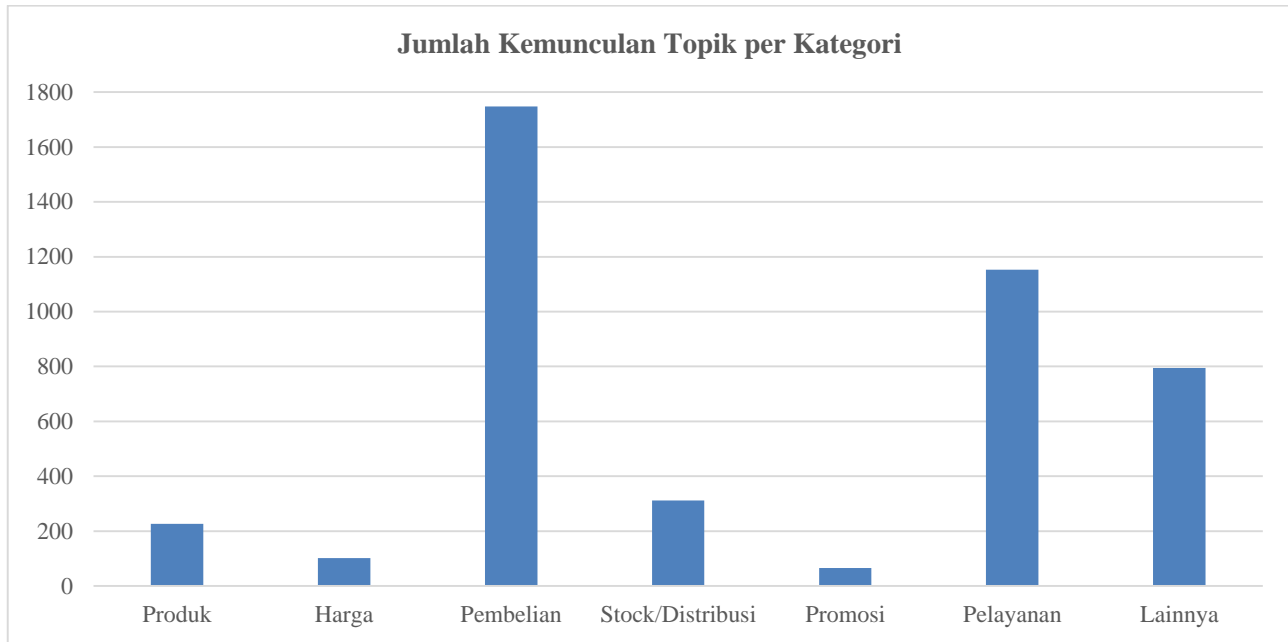


Gambar 6. 15. Distribusi Tiap Topik

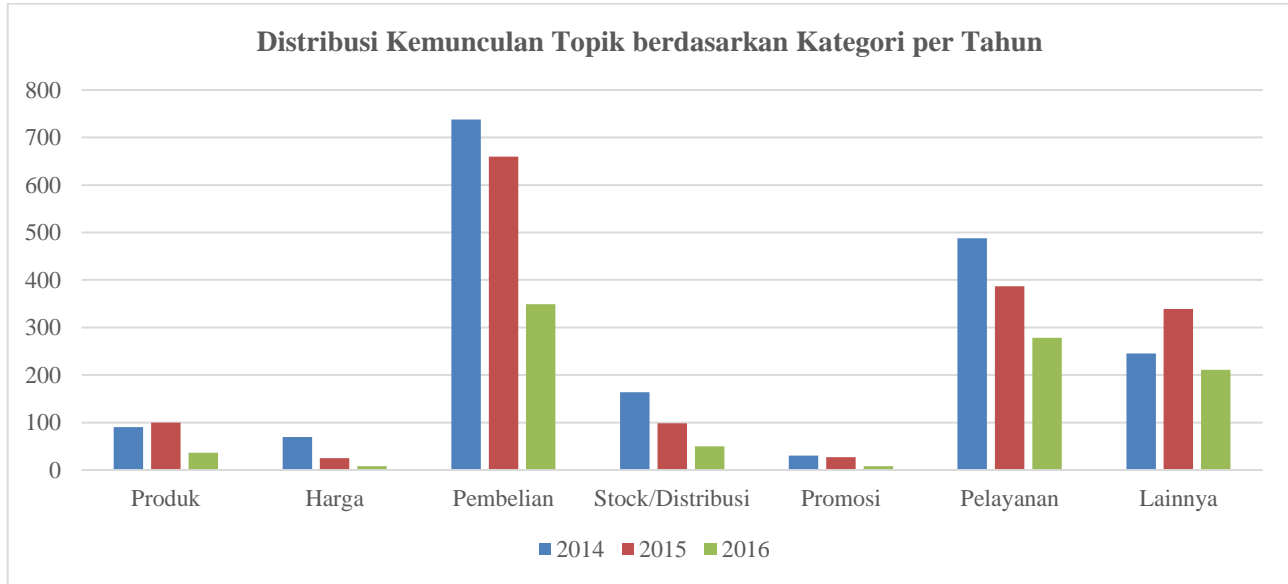
Setiap topik digolongkan ke dalam kategori topik tertentu. Gambar 6.16. menunjukkan bahwa distribusi dokumen pada setiap kategori memiliki jumlah paling besar pada kategori topik pembelian dengan jumlah kemunculan pada 1748 dokumen. Sedangkan kategori topik yang memiliki jumlah kemunculan dokumen yang besar selanjutnya adalah kategori pelayanan dengan jumlah 1153. Kemudian dilanjutkan dengan kategori lainnya yaitu dengan jumlah kemunculan dokumen sebanyak 795.

Gambar 6.17. menunjukkan distribusi dokumen pada setiap kategori dan tahun yang berbeda. Pada tahun 2014, kategori topik yang paling sering dibahas oleh pelanggan adalah kategori pembelian dengan jumlah kemunculan dokumen sebanyak 738 dokumen. Sama halnya dengan tahun 2015, kategori yang paling sering dibahas oleh pelanggan adalah mengenai pembelian. Namun, terjadi penurunan jumlah kemunculan dokumen yaitu sebanyak 600 dokumen. Hal ini dikarenakan bahwa jumlah suara pelanggan yang juga menurun pada tahun 2015. Sedangkan pada tahun 2016, kategori pembelian juga memiliki jumlah kemunculan paling banyak yaitu sejumlah 349 dokumen.

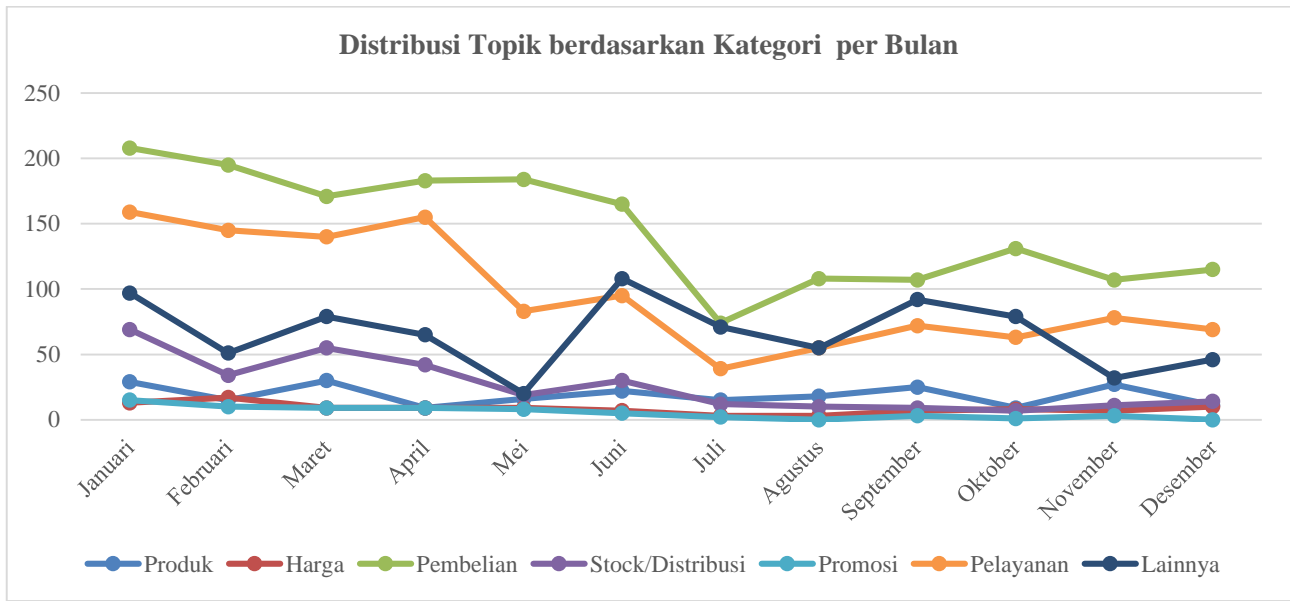
Gambar 6.18 menunjukkan jumlah total tiap kategori topik dilihat berdasarkan bulan. Dalam grafik, terlihat bahwa, topik yang paling sering dibicarakan tiap bulan adalah kategori pembelian. Sedangkan yang paling jarang dibahas oleh pelanggan adalah kategori promosi dan harga dengan jumlah kemunculan dokumen rata-rata berkisar antara 0-20 dokumen setiap bulan pada keseluruhan tahun 2014, 2015, dan 2016. Pada bulan juli terdapat penurunan yang signifikan terhadap jumlah dokumen yang ada. Hal ini disebabkan karena data dalam penelitian ini yang digunakan adalah data pada bulan Januari 2014 hingga Juni 2016.



Gambar 6. 16. Distribusi Topik per Kategori



Gambar 6. 17. Distribusi Topik Tiap Kategori per Tahun



Gambar 6. 18. Total Distribusi Topik Tiap Kategori per Bulan

6.6. Pengujian Model pada Aplikasi

Pengujian dalam penelitian ini menggunakan data testing sejumlah 227 data, yaitu data suara pelanggan pada bulan Juli dan Agustus 2016. Data dilabelkan berdasarkan topik yang telah teridentifikasi sebelumnya. Dari hasil pengujian, didapatkan 19 topik yang sesuai dengan 35 topik yang telah ada. Dari total 227 data tersebut, data yang tepat terlabelkan adalah sejumlah 190 data, sedangkan data yang memiliki perbedaan label dari yang seharusnya adalah 37 data. Sehingga, jika diukur dengan menggunakan prosentase, maka akurasi data adalah :

$$akurasi = \frac{\text{jumlah data tepat berlabel}}{\text{total seluruh data}}$$

sehingga kemudian didapatkan akurasi = $190/227$ adalah 83.7%. Sedangkan tingkat error atau error rate dari penelitian ini adalah $37/227$ atau sama dengan 16.3%. Hal ini dipengaruhi oleh beberapa faktor. Faktor pertama yang paling mempengaruhi tingkat akurasi adalah distribusi topik dalam dokumen. Suatu dokumen dapat memiliki kemungkinan untuk masuk ke dalam beberapa topik. Hal tersebut akan membuat nilai probabilitas suatu dokumen menjadi rata, sehingga dokumen dapat dikatakan sebagai dokumen yang memiliki mutual exclusive ke dalam topik lain. Faktor kedua yang mempengaruhi hasil pengujian adalah penamaan topik berdasarkan kata-kata yang terkandung di dalamnya. Semakin erat nama topik dengan kata-kata yang terkandung didalam topik, akan memberikan nilai yang semakin baik pula. Hasil dari pengujian terhadap model dalam aplikasi ditunjukkan pada Lampiran-B.

BAB VII

KESIMPULAN DAN SARAN

Bab ini menjelaskan hasil dari implementasi perancangan studi kasus atau hasil dari proses pelaksanaan penelitian. Hasil yang akan dijabarkan adalah hasil wawancara dengan narasumber dan hasil dari analisis data dari arsip dokumen milik organisasi. Selain itu, akan dijelaskan juga mengenai hambatan dan rintangan dalam proses pelaksanaan penelitian.

7.1.Kesimpulan

Berdasarkan hasil dari Tugas Akhir yang dikerjakan oleh penulis, didapatkan kesimpulan yang juga ditarik berdasar dari tujuan awal pengerjaan tugas akhir ini sebagai berikut :

1. Analisis topic modelling menggunakan Latent Dirichlet Allocation pada penelitian ini menggunakan 4400 dokumen Suara Pelanggan Bulan Januari 2014-Juni 2016 dengan melakukan stemming disertai pengecekan yang selanjutnya menghasilkan model LDA dengan iterasi sejumlah 20 kali dan jumlah topik=35. Dari 35 topik yang dihasilkan, topik kemudian disesuaikan dengan panduan identifikasi permasalahan pada perusahaan sehingga menjadi 26 topik dengan 7 kategori. Dokumen kemudian diklasifikasikan berdasarkan topik yang terbentuk. Hasil dari pengujian dokumen terhadap topik menghasilkan ketepatan sebesar 83.7% yaitu 190 dokumen dari 227 dokumen.
2. Dokumen yang telah diklasifikasikan berdasarkan topik yang telah diidentifikasi, divisualisasikan ke dalam bar chart untuk mengetahui jumlah kategori topik yang sering dibahas oleh pelanggan pada tahun 2016. Kategori produk memiliki jumlah kemunculan topik sebanyak 36 dokumen,

kategori harga memiliki jumlah kemunculan topik sebanyak 8 dokumen, kategori pembelian memiliki jumlah kemunculan topik sebanyak 349 dokumen, kategori stock/distribusi memiliki jumlah kemunculan topik sebanyak 50 dokumen, kategori promosi memiliki jumlah kemunculan topik sebanyak 8 dokumen, kategori pelayanan memiliki jumlah kemunculan topik sebanyak 278 dokumen dan kategori lainnya memiliki jumlah kemunculan topik sebanyak 211 dokumen.

7.2.Saran

Saran dari pengerjaan tugas akhir ini adalah dengan memperbanyak parameter pemilihan jumlah topik yang dihasilkan. Dalam penelitian ini parameter pemilihan jumlah topik dirasa kurang menunjukkan keseluruhan topik seperti yang diharapkan oleh perusahaan, namun jika jumlah topik diperbanyak, maka masing-masing topik memiliki kemiripan antara satu topik dengan topik lain. Selain itu, pada tahap pra pemrosesan, perlu ditambah dengan pendefinisian term, seperti kata yang dipisah dan kata yang digabung yang seharusnya memiliki arti yang sama. Selain term, penelitian selanjutnya juga dapat mempertimbangkan field data yang digunakan untuk pembuatan model. Dalam penelitian ini, 4 field yang dipilih digabungkan untuk mendapatkan topik dalam dokumen. Field mempengaruhi isi kata-kata dalam dokumen. Oleh karena itu, perlu dilakukan penyeleksian field yang paling penting dan menggambarkan topik untuk dianalisis. Serta penelitian selanjutnya, perlu mempertimbangkan aspek kestabilan suatu model ketika dijalankan. Aspek kestabilan mempengaruhi hasil topik yang muncul pada setiap kali model dijalankan.

REFERENSI

- [1] PT. Petrokimia Gresik, 2015. [Online]. Available: <http://www.petrokimia-gresik.com/>.
- [2] A. Ghaziana, M. H. Hossainib dan H. Farsijanic, “The effect of customer relationship management and its significant,” 2016.
- [3] J.-F. Yeh, Y.-S. Tan dan C.-H. Lee, “Topic detection and tracking for conversational content by using conceptual dynamic latent Dirichlet Allocation,” *Neurocomputing*, 2016.
- [4] M. H. Basri, “IDENTIFIKASI TOPIK INFORMASI PUBLIK MEDIA SOSIAL DI KOTA SURABAYA BERDASARKAN KLASTERISASI TEKS PADA TWITTER DENGAN MENGGUNAKAN ALGORITMA K-MEANS,” 2016.
- [5] E. Gallinucci, M. Golfarelli dan S. Rizzi, “Advanced topic modeling for social business intelligence,” 2014.
- [6] A. Gaur, “TOPIC MODELS AS A NOVEL APPROACH TO IDENTIFY THEMES IN CONTENT ANALYSIS: THE EXAMPLE OF ORGANIZATIONAL RESEARCH METHODS”.
- [7] D. M. Blei, “Latent Dirichlet Allocation,” *Machine Learning Research* 3, pp. 933-1022, 2003.
- [8] M. R. Brett, “Journal of Digital Humanities,” Desember 2012. [Online]. Available: <http://journalofdigitalhumanities.org/2-1/topic-modeling-a-basic-introduction-by-megan-r-brett/>.
- [9] D. M. Blei, “Probabilistic Topic Model,” *communications of the acm*, vol. 55, 2012.

- [10] E. Chen, "Introduction to Latent Dirichlet Allocation," 2011. [Online]. Available: <http://blog.echen.me/2011/08/22/introduction-to-latent-dirichlet-allocation/>.
- [11] A. Knispelis, "LDA Topic Models," Youtube.
- [12] S. Moro, P. Cortez dan P. Rita, "Business intelligence in banking: A literature analysis from 2002 to 2013," *Elsevier*, 2014.
- [13] J. C. Campbell, A. Hindle dan a. E. Stroulia, "Latent Dirichlet Allocation: Extracting Topics," 2014.
- [14] R. Tavalai, "Customer-oriented principles in Islam," *police bimonthly human development* , no. 17, 2008.
- [15] K. Feizi, "Customer Relationship Management and Customer Satisfaction : An Analysis Based on Quality of Service (Srkval) (Company E Corp)," *improvement and change management studies*, vol. 57, no. 105-132, 2008.
- [16] N. Morgan, B. Clark dan R. Gooner, "Marketing productivity, marketing audits and systems for marketing performance assesment: integrating multiple perspective," *Journal of Business Research*, vol. 55, 2002.

BIODATA PENULIS



Penulis lahir di Tuban, 31 Agustus 1995. Karena nama yang terlalu mainstream, penulis memiliki nama panggilan yang unik, diantaranya “Aree”, “Arik”, atau “Ari”. Penulis menempuh pendidikan formal di SD Negeri Sokosari 1, SMP Negeri 1 Bojonegoro, SMA Negeri 1 Bojonegoro. Setelah selesai menempuh pendidikan SMA, penulis melanjutkan pendidikan di Jurusan Sistem Informasi FTIf-ITS dan menjadi mahasiswa Sistem Informasi tahun angkatan 2013. Penulis aktif menjadi mahasiswa yang mengikuti banyak kegiatan baik akademik maupun organisasi selama perkuliahan. Hal ini dibuktikan dengan menjadi staff Himpunan Mahasiswa Sistem Informasi ITS periode 2014/2015 dan Kepala Divisi Departemen Sosial Masyarakat HMSI Muda Berkarya 2015/2016, staff Kementerian Sosial Masyarakat BEM ITS 2014/2015, Pemandu FTIf, serta mengikuti berbagai kepanitian dan pelatihan. Semoga pengalaman penulis serta penulisan Tugas Akhir ini mampu memberikan kontribusi positif bagi semua pihak terkait.

Halaman sengaja dikosongkan

LAMPIRAN A

Evaluasi Jumlah Iterasi

Tabel A. 1. Perplexity Percobaan 1 Tanpa Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	103.1	108.2	112.5
2	69.4	63.3	57.6
3	64.5	57.6	52.8
4	62.1	54.9	50.8
5	60.8	53.6	49.7
6	59.8	52.9	49.1
7	59.1	52.4	48.7
8	58.5	52.1	48.4
9	58.1	51.8	48.2
10	57.9	51.6	48.1
11	57.7	51.5	48
12	57.5	51.4	47.9
13	57.3	51.3	47.8
14	57.1	51.2	47.7
15	57	51.1	47.6
16	56.9	51.1	47.6
17	56.8	51	47.5
18	56.7	51	47.5
19	56.6	50.9	47.4
20	56.5	50.9	47.4
21	56.4	50.8	47.4
22	56.3	50.8	47.3
23	56.3	50.8	47.3

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
24	56.2	50.7	47.3
25	56.2	50.7	47.2
26	56.1	50.7	47.2
27	56.1	50.7	47.2
28	56	50.6	47.2
29	56	50.6	47.2
30	55.9	50.6	47.1

Tabel A. 2. Perplexity Percobaan 2 Tanpa Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	94.1	102	108.6
2	68.3	60.5	55.6
3	59.3	56.3	51.7
4	57.2	54.1	50
5	56	52.6	49.1
6	55.2	51.6	48.6
7	54.7	50.9	48.2
8	54.3	50.5	47.9
9	53.9	50.2	47.7
10	53.7	50	47.5
11	53.4	49.9	47.4
12	53.2	49.7	47.3
13	53	49.6	47.2
14	52.9	49.6	47.1
15	52.7	49.5	47.1
16	52.6	49.4	47
17	52.4	49.4	47

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
18	52.3	49.3	46.9
19	52.2	49.3	46.9
20	52	49.2	46.8
21	51.8	49.2	46.8
22	51.8	49.2	46.8
23	51.7	49.2	46.7
24	51.7	49.1	46.7
25	51.6	49.1	46.7
26	51.6	49.1	46.7
27	51.6	49.1	46.6
28	51.5	49	46.6
29	51.5	49	46.6
30	51.5	49	46.6

Tabel A. 3. Perplexity Percobaan 3 Tanpa Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	100	100.1	114.3
2	67.1	57.3	57.5
3	60.7	53.6	52.3
4	58.2	52.3	50.1
5	57	51.5	49.1
6	56.3	51.1	48.5
7	55.8	50.8	48.1
8	55.4	50.6	47.8
9	55.1	50.4	47.6
10	54.9	50.3	47.4
11	54.8	50.2	47.3

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
12	54.6	50.1	47.2
13	54.5	50	47.1
14	54.4	49.9	47
15	54.3	49.9	47
16	54.2	49.8	46.9
17	54.1	49.8	46.8
18	54.1	49.7	46.8
19	54	49.7	46.8
20	54	49.6	46.7
21	53.9	49.6	46.7
22	53.9	49.6	46.7
23	53.8	49.5	46.6
24	53.8	49.5	46.6
25	53.7	49.5	46.6
26	53.7	49.4	46.5
27	53.7	49.4	46.5
28	53.6	49.4	46.5
29	53.6	49.4	46.5
30	53.6	49.4	46.5

Tabel A. 4. Perplexity Percobaan 1 dengan Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	73.3	78.1	85.7
2	51.6	49.6	47.8
3	48.8	46.8	44.2
4	47.6	45.5	42.6
5	47	44.9	41.7

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
6	46.6	44.5	41.2
7	46.3	44.2	40.8
8	46.1	44	40.5
9	45.9	43.8	40.3
10	45.8	43.7	40.1
11	45.7	43.6	40
12	45.6	43.5	39.8
13	45.5	43.4	39.7
14	45.4	43.3	39.7
15	45.3	43.3	39.6
16	45.2	43.2	39.5
17	45.1	43.2	39.5
18	45.1	43.1	39.4
19	45	43.1	39.4
20	44.9	43	39.3
21	44.9	43	39.3
22	44.9	43	39.3
23	44.8	42.9	39.2
24	44.8	42.9	39.2
25	44.8	42.9	39.2
26	44.7	42.9	39.2
27	44.7	42.8	39.1
28	44.7	42.8	39.1
29	44.7	42.8	39.1
30	44.6	42.8	39.1

Tabel A. 5. Perplexity Percobaan 2 dengan Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	78.4	75.4	89.1
2	54.4	47.5	49.1
3	51.3	44.7	45.2
4	49.9	43.5	43.5
5	49.2	42.8	42.6
6	48.8	42.4	42
7	48.5	42.1	41.6
8	48.3	41.9	41.4
9	48.1	41.8	41.2
10	48.1	41.6	41
11	48	41.5	40.9
12	47.8	41.4	40.8
13	47.7	41.4	40.7
14	47.5	41.3	40.6
15	47.4	41.2	40.5
16	47.3	41.2	40.5
17	47.2	41.1	40.4
18	47.1	41.1	40.4
19	47.1	41.1	40.3
20	47	41	40.3
21	47	41	40.3
22	46.9	41	40.2
23	46.9	41	40.2
24	46.8	40.9	40.1
25	46.8	40.9	40.1
26	46.7	40.9	40.1
27	46.7	40.9	40.1

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
28	46.7	40.9	40
29	46.7	40.8	40
30	46.6	40.8	40

Tabel A. 6. Perplexity Percobaan 3 dengan Stemming

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	76	80.3	82.9
2	50.4	50.2	46.8
3	47.2	46.7	43.7
4	45.8	44.9	42.3
5	45.1	43.9	41.4
6	44.6	43.3	40.9
7	44.3	42.9	40.5
8	44.1	42.6	40.3
9	44	42.4	40
10	43.9	42.2	39.9
11	43.8	42.1	39.7
12	43.7	42	39.6
13	43.6	41.9	39.5
14	43.5	41.9	39.4
15	43.5	41.8	39.3
16	43.5	41.7	39.3
17	43.4	41.7	39.2
18	43.4	41.6	39.1
19	43.3	41.6	39.1
20	43.3	41.6	39
21	43.3	41.5	39

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
22	43.2	41.5	39
23	43.2	41.5	38.9
24	43.2	41.5	38.9
25	43.2	41.4	38.9
26	43.1	41.4	38.8
27	43.1	41.4	38.8
28	43.1	41.4	38.8
29	43.1	41.4	38.8
30	43.1	41.4	38.7

Tabel A. 7. Perplexity Percobaan 1 dengan Stemming dan Pengecekan

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	77.3	78.7	84.3
2	54.5	48.2	45.7
3	51	44.2	42.3
4	49.2	41.6	40.8
5	48.4	40.1	39.9
6	47.9	39.5	39.3
7	47.5	39.1	38.9
8	47.3	38.9	38.6
9	46.9	38.7	38.4
10	46.5	38.6	38.2
11	46.3	38.5	38
12	46.2	38.4	37.9
13	46.1	38.3	37.8
14	46	38.2	37.8
15	46	38.1	37.7

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
16	45.9	38	37.6
17	45.8	38	37.6
18	45.8	37.9	37.6
19	45.7	37.9	37.5
20	45.7	37.8	37.5
21	45.6	37.8	37.5
22	45.6	37.8	37.4
23	45.5	37.7	37.4
24	45.5	37.7	37.4
25	45.5	37.7	37.4
26	45.4	37.7	37.3
27	45.4	37.6	37.3
28	45.4	37.6	37.3
29	45.4	37.6	37.3
30	45.3	37.6	37.3

Tabel A. 8. Perplexity Percobaan 2 dengan Stemming dan Pengecekan

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	73.7	81.6	89.7
2	49.4	50.3	48
3	46.7	47.1	44
4	45.6	45.4	42.2
5	44.8	44.4	41
6	44.3	43.8	40.3
7	44	43.3	39.8
8	43.8	43	39.5
9	43.6	42.7	39.3

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
10	43.5	42.4	39.1
11	43.4	42.3	39
12	43.3	42.1	38.9
13	43.1	41.9	38.8
14	43	41.8	38.7
15	43	41.7	38.6
16	42.9	41.6	38.6
17	42.8	41.6	38.5
18	42.8	41.5	38.5
19	42.7	41.4	38.4
20	42.7	41.4	38.4
21	42.6	41.3	38.4
22	42.6	41.3	38.3
23	42.6	41.3	38.3
24	42.5	41.2	38.3
25	42.5	41.2	38.3
26	42.5	41.2	38.2
27	42.4	41.1	38.2
28	42.4	41.1	38.2
29	42.4	41.1	38.2
30	42.4	41.1	38.1

Tabel A. 9. Perplexity Percobaan 3 dengan Stemming dan Pengecekan

Jumlah Iterasi	5 Topik	10 Topik	15 Topik
1	74.6	75.8	83.8
2	49.5	46.7	44.6
3	46.6	44.3	41.3

4	45.2	43	40
5	44.5	42.1	39.3
6	44	41.6	38.9
7	43.7	41.2	38.7
8	43.5	41	38.5
9	43.3	40.8	38.3
10	43.2	40.6	38.2
11	43	40.5	38.1
12	42.9	40.4	38
13	42.8	40.3	37.9
14	42.7	40.2	37.8
15	42.6	40.2	37.8
16	42.3	40.1	37.7
17	42.1	40.1	37.7
18	41.8	40	37.6
19	41.6	40	37.6
20	41.4	39.9	37.5
21	41.3	39.9	37.5
22	41.1	39.9	37.5
23	41.1	39.8	37.5
24	40.8	39.8	37.4
25	40.6	39.8	37.4
26	40.6	39.8	37.4
27	40.6	39.7	37.4
28	40.5	39.7	37.4
29	40.5	39.7	37.3
30	40.5	39.7	37.3

Tabel A. 10. Hasil Uji Coba Model

Jumlah Topik	Percobaan										Mean	Stdev
	1	2	3	4	5	6	7	8	9	10		
3	48.7	50.7	48.5	48.4	50.1	49.2	51.9	50.1	48.1	48.7	49.44	1.221293
4	44.4	46.1	46.7	45.4	46.8	47.2	44.1	45.9	43	43.8	45.34	1.437745
5	43.8	44.1	45.8	45	41.2	41.7	43.1	44	42.7	41.3	43.27	1.559238
6	40.7	41.6	45.2	42.4	42.3	39.5	39.8	40.7	42.4	42	41.66	1.637206
7	42.8	42.2	41.8	38.4	41	41.3	41.5	41.1	40.7	41	41.18	1.164092
8	40	38.8	40.6	41.9	40.5	41.8	40.5	39.8	40.6	41.3	40.58	0.935474
9	40.2	38.9	38.5	38.7	40.4	38.7	40.6	39.7	39.3	39.4	39.44	0.760409
10	37.9	41.1	40.3	40.9	39.8	40	41.6	38.3	40.5	39	39.94	1.212161
11	41.7	38.5	39.1	37.1	40.5	39.7	38.9	38.1	39.9	38.6	39.21	1.301666
12	38.4	39.3	39.7	37.4	37.7	38.6	39.6	39.1	37.9	38.3	38.6	0.804156
13	37.7	40.5	38.7	37.5	37.6	38.3	38.2	39	39.4	40	38.69	1.031127
14	38.6	39.2	37.4	38.6	37.4	38.1	38	40.3	37.6	39.5	38.47	0.964999

Jumlah Topik	Percobaan										Mean	Stdev
	1	2	3	4	5	6	7	8	9	10		
15	37.4	36.9	37.5	38.3	37.8	36.1	38.3	39.2	37.7	37.8	37.7	0.837987
16	36.9	36.8	36.9	37.6	37.1	38.6	38.7	36.9	38.4	38.2	37.61	0.786624
17	37.2	38.6	36.7	37.1	36.5	38.7	37.6	36.5	37.7	36.7	37.33	0.812472
18	37.9	37	40.1	38.3	37.4	37.2	35.7	37	36.8	37.1	37.45	1.155903
19	36.2	36.8	36.2	37	37	35.4	37.8	36.3	36.7	35.9	36.53	0.675031
20	36.4	37.5	36	36.3	35.7	37.4	37.7	36.2	37.2	36.2	36.66	0.715231
21	36.3	36.4	35.8	37.2	36.6	35.6	36.2	37.4	35.5	36.4	36.34	0.623966
22	37.6	36.8	38.2	37	35.4	36	36.7	36.4	35.4	35.2	36.47	0.991127
23	37.4	36.6	35.8	35.5	36.5	35.8	35.7	35.9	35.5	35.5	36.02	0.623253
24	36.5	35.8	36.5	35.6	34.5	35.8	35.2	35.3	37.2	35	35.74	0.808565
25	35.3	35.2	36.1	36	35.4	35.7	36.3	37.1	35.3	36.3	35.87	0.605622
26	35.6	35.6	35.9	35.2	36	36	36	36.2	36	34.8	35.73	0.43729
27	35.6	36.3	36.3	35.6	36.4	36.2	36.2	34.9	34.7	34	35.62	0.829726

Jumlah Topik	Percobaan										Mean	Stdev
	1	2	3	4	5	6	7	8	9	10		
28	34.8	35.1	35.1	35.3	36.2	33.7	35.5	36.1	35.5	34.6	35.19	0.729459
29	35.4	35.5	35.2	36.2	35.8	36	35.7	34.8	35.7	35.8	35.61	0.404008
30	36.3	35.8	35.6	35.1	35.2	35	34.7	36	35.6	34.5	35.38	0.576965
31	36.1	36.9	36.1	36.6	37.4	36.3	37.7	36.8	36.8	35.6	36.63	0.632543
32	35	37.4	36.6	36.8	37.8	36.2	38.5	37	36.5	36.6	36.84	0.945398
33	36.9	36.3	36.4	36.5	37.1	37	36	36	35.9	36.4	36.45	0.430116
34	36.7	35.6	36.5	36.8	36.3	36.2	35.9	35	37	35.4	36.14	0.653537
35	35.1	35.8	34.9	34.9	35.1	34.5	35.1	34.7	33.9	35.2	34.92	0.496208
36	33.3	34.8	35	35.5	35	34.8	35.2	35.3	34.8	34.7	34.84	0.598517
37	34.8	35.2	34.4	35.5	34.4	35	35.3	35	33.9	35.2	34.87	0.496767
38	33.8	35.1	35.1	34.4	34.5	36.1	34.5	34.1	35.6	35.3	34.85	0.712195
39	34.4	33.7	34	34.4	34.9	35.3	35.3	33.7	35	34.7	34.54	0.602218
40	34.3	33.9	35.1	34.8	34.1	35.1	34.1	35	35.3	34.2	34.59	0.519508

Jumlah Topik	Percobaan										Mean	Stdev
	1	2	3	4	5	6	7	8	9	10		
41	34.8	37.7	36.2	35.6	36.5	35.1	35.8	34.6	35.8	36.2	35.83	0.9056
42	36.1	35.9	36.3	36.5	36.7	35.6	35.4	36.2	36.6	36.2	36.15	0.419656
43	36.4	35.8	35.6	35.5	35.5	36.1	36.1	35.7	36.3	35.9	35.89	0.324722
44	35.5	35.3	35.4	35.9	35.8	34.5	34.8	34.4	35.7	34	35.13	0.660051
45	34.4	34.8	34	34.7	34.3	35.1	33.5	33.9	34.6	33.9	34.32	0.493964
46	34.5	34	35.6	34.1	34.4	34.8	35	35.9	34.1	34.3	34.67	0.653282
47	34.5	34.9	35	34.8	34.1	34.3	33.8	34	33.5	35.2	34.41	0.562633
48	34.5	34.1	34.2	34.5	34.6	34.1	34	34.9	33.8	34.1	34.28	0.332666
49	35	34.6	36.1	34.2	35	35.7	36.8	35.4	35.3	35.7	35.38	0.748034
50	36.1	35.8	35.3	35.4	37.2	35.9	35.4	36	35.6	35.3	35.8	0.573488

Halaman sengaja dikosongkan

LAMPIRAN B

Tabel B. 1. Hasil Pengujian Data

Dokumen Ke	Topik	Kesesuaian
0	7	Ya
1	3	Ya
2	3	Ya
3	33	Ya
4	3	Tidak
5	3	Ya
6	7	Ya
7	14	Ya
8	20	Ya
9	12	Ya
10	14	Ya
11	19	Ya
12	19	Ya
13	14	Ya
14	19	Ya
15	12	Tidak
16	33	Ya
17	20	Tidak
18	19	Ya
19	14	Ya
20	19	Ya
21	19	Ya
22	19	Ya
23	19	Ya
24	3	Tidak
25	3	Tidak

Dokumen Ke	Topik	Kesesuaian
26	3	Ya
27	19	Ya
28	2	Ya
29	28	Ya
30	28	Ya
31	20	Ya
32	12	Ya
33	24	Ya
34	16	Ya
35	24	Ya
36	24	Ya
37	24	Ya
38	23	Ya
39	14	Ya
40	3	Tidak
41	33	Ya
42	3	Ya
43	3	Ya
44	18	Ya
45	14	Ya
46	12	Ya
47	7	Ya
48	13	Ya
49	12	Ya
50	12	Ya
51	18	Ya
52	3	Ya
53	3	Ya
54	20	Ya

Dokumen Ke	Topik	Kesesuaian
55	2	Ya
56	28	Ya
57	20	Ya
58	14	Ya
59	3	Ya
60	20	Ya
61	3	Tidak
62	3	Ya
63	2	Ya
64	20	Tidak
65	2	Ya
66	16	Ya
67	3	Ya
68	3	Ya
69	24	Ya
70	3	Tidak
71	20	Ya
72	12	Ya
73	3	Tidak
74	16	Ya
75	16	Ya
76	16	Ya
77	33	Ya
78	3	Tidak
79	19	Ya
80	14	Ya
81	14	Ya
82	24	Ya
83	24	Ya

Dokumen Ke	Topik	Kesesuaian
84	24	Ya
85	14	Ya
86	19	Ya
87	24	Ya
88	7	Ya
89	4	Ya
90	28	Ya
91	4	Ya
92	13	Tidak
93	14	Ya
94	14	Ya
95	24	Ya
96	12	Ya
97	12	Ya
98	12	Ya
99	7	Ya
100	7	Ya
101	12	Ya
102	20	Ya
103	3	Ya
104	3	Tidak
105	2	Tidak
106	12	Ya
107	14	Tidak
108	12	Ya
109	19	Ya
110	12	Ya
111	12	Ya
112	14	Ya

Dokumen Ke	Topik	Kesesuaian
113	13	Ya
114	2	Ya
115	2	Tidak
116	12	Ya
117	12	Ya
118	7	Ya
119	33	Ya
120	28	Ya
121	24	Ya
122	4	Ya
123	12	Ya
124	12	Ya
125	12	Ya
126	7	Ya
127	2	Ya
128	2	Ya
129	12	Ya
130	7	Ya
131	4	Ya
132	3	Ya
133	7	Ya
134	7	Ya
135	2	Ya
136	12	Ya
137	12	Ya
138	3	Tidak
139	2	Ya
140	12	Ya
141	7	Ya

Dokumen Ke	Topik	Kesesuaian
142	7	Ya
143	20	Tidak
144	2	Ya
145	7	Ya
146	7	Ya
147	2	Ya
148	3	Tidak
149	7	Ya
150	4	Tidak
151	3	Tidak
152	7	Ya
153	7	Tidak
154	4	Ya
155	33	Ya
156	16	Ya
157	7	Tidak
158	7	Ya
159	7	Ya
160	7	Tidak
161	2	Ya
162	12	Ya
163	4	Ya
164	2	Ya
165	4	Ya
166	28	Ya
167	2	Ya
168	3	Tidak
169	4	Ya
170	4	Ya

Dokumen Ke	Topik	Kesesuaian
171	20	Tidak
172	20	Tidak
173	7	Ya
174	28	Ya
175	20	Ya
176	20	Ya
177	19	Ya
178	19	Ya
179	19	Ya
180	19	Ya
181	19	Ya
182	24	Ya
183	24	Ya
184	14	Ya
185	19	Ya
186	19	Ya
187	0	Ya
188	17	Tidak
189	19	Ya
190	17	Tidak
191	17	Tidak
192	17	Tidak
193	17	Tidak
194	17	Tidak
195	24	Ya
196	24	Ya
197	24	Ya
198	24	Ya
199	24	Ya

Dokumen Ke	Topik	Kesesuaian
200	24	Ya
201	14	Ya
202	24	Ya
203	24	Ya
204	14	Ya
205	23	Ya
206	14	Ya
207	0	Ya
208	14	Ya
209	23	Ya
210	24	Ya
211	14	Ya
212	19	Ya
213	19	Ya
214	0	Ya
215	0	Ya
216	14	Ya
217	14	Ya
218	24	Ya
219	14	Ya
220	24	Ya
221	24	Ya
222	0	Ya
223	17	Tidak
224	25	Tidak
225	15	Tidak
226	1	Tidak

Tabel B. 2. Hasil Klasifikasi Jumlah Dokumen berdasarkan ID Topik

ID Topik	Jumlah Kemunculan
0	102
1	54
2	142
3	518
4	507
5	65
6	9
7	567
8	102
9	25
10	5
11	39
12	156
13	71
14	129
15	10
16	256
17	39
18	192
19	182
20	229
21	19
22	5
23	93
24	159
25	48
26	2
27	34

ID Topik	Jumlah Kemunculan
28	188
29	24
30	23
31	42
32	59
33	305
34	1

Tabel B. 3. Hasil Klasifikasi Data berdasarkan Topik

Jumlah	Topik
2	Alamat Distributor
282	Dosis Pemupukan
102	Harga Pupuk
241	Kabar Positif Perusahaan
48	Kantong Pupuk
350	Kegiatan Perusahaan
71	Kemasan Produk
117	Kualitas Perusahaan
229	Lowongan Pekerjaan
54	Manfaat Produk
674	Pembelian Produk Non Pupuk
507	Pembelian Produk Subsidi
567	Pembelian Pupuk Non Subsidi
5	Penangan Masalah
34	Penggunaan Produk
142	Penjelasan Seputar Perusahaan
23	Perjanjian
39	Permasalahan Distribusi Produk
59	Permohonan Kunjungan dan

Jumlah	Topik
	Demplot
39	Produksi
5	Sosialisasi Berimbang
234	Stock Produk
188	Syarat Menjadi Distributor
305	Syarat menjadi kios resmi
65	Tabloid Sahabat Petani

Tabel B. 4. Hasil Klasifikasi Data berdasarkan Kategori

Jumlah	Kategori
226	Produk
102	Harga
1748	Pembelian
312	Stock/Distribusi
65	Promosi
1153	Pelayanan
795	Lainnya

Tabel B. 5. Hasil Klasifikasi Data berdasarkan Kategori Tiap Tahun

Kategori	Tahun		
	2014	2015	2016
Produk	90	100	36
Harga	69	25	8
Pembelian	738	660	349
Stock/Distribusi	164	98	50
Promosi	30	27	8
Pelayanan	488	387	278
Lainnya	245	339	211

Tabel B. 6. Hasil Klasifikasi Data per Kategori Berdasarkan Bulan

Kategori	Bulan											
	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Ags	Sep	Okt	Nov	Des
Produk	29	15	30	9	16	22	15	18	25	9	27	11
Harga	13	17	9	9	9	7	3	3	7	8	7	10
Pembelian	208	195	171	183	184	165	74	108	107	131	107	115
Stock/Distribusi	69	34	55	42	19	30	12	10	9	7	11	14
Promosi	15	10	9	9	8	5	2	0	3	1	3	0
Pelayanan	159	145	140	155	83	95	39	55	72	63	78	69
Lainnya	97	51	79	65	20	108	71	55	92	79	32	46

Data Kategorisasi Topik Berdasarkan Topik yang teridentifikasi

Produk

Manfaat Produk [1], Kemasan Produk [13], Warna Pupuk [21], Kantong Pupuk [25], Penggunaan produk [27]

Harga

Harga pupuk [8]

Pembelian

Info pembelian produk : Pupuk Subsidi [4], Pupuk Non Subsidi [7], Non Pupuk [3, 12]

Stock/Distribusi

Produksi [11], Permasalahan Distribusi Produk [17], Stock Produk [18, 31]

Promosi

Tabloid Sahabat Petani [5]

Pelayanan

Penjelasan non produk : penjelasan seputar perusahaan [2], sosialisasi berimbang [22], alamat distributor [26]

Lowongan pekerjaan [20]

Syarat menjadi distributor [28]

Syarat menjadi kios resmi [33]

Dosis Pemupukan [9, 16, 34]

Lainnya

Kerjasama : kualitas perusahaan [23, 29], perjanjian [30]

Permohonan kunjungan dan demplot [32]

Pemberitaan : Kabar positif perusahaan [0, 14, 15], kegiatan perusahaan [6, 19, 24], penanganan masalah [10]